Seven Discourses

on

the Ontology of Natural Languages

Paul M. Postal

Contents

Preface

It seems uncontroversial that there are natural languages and that there is a field, linguistics, which studies them, or should. Granting the existence of natural languages raises questions about their ontology and that of their elements, sentences. What kind of things are they? Where in the overall universe of types of things, protons, trees, chipmunks, integers, logic principles, books, dreams do they lie?

To a certain extent, to be sure, various kinds of linguistic research are possible without explicit answers to such questions but in general, this is far from an optimal state of affairs. In some cases it is, I suggest, downrightly harmful.

If this point of view is correct, consideration of ontological questions, usually backgrounded in most linguistic research, is a worthwhile enterprise. In any event, this is the kind of thing considered below. Readers will have to decide whether there is any merit in the project itself and any value in this particular pursuit of it.

Part I

Part I originated as the opening remarks at the International Workshop 'Foundations of Linguistics', Technical University Braunschweig, Germany, presented (via video) on June 27 2015, organized by Christina Behme and Martin Neef. This may have been the first Earth conference devoted exclusively to the ontology of natural language. The original remarks have been slightly modified to take account of the present written context. A version of Part I appears under the name 'The ontology of natural language' in Behme and Neef (2018).

In the award-winning 2002 *Cambridge Grammar of the English Language*, the editors say:

(1) Huddleston and Pullum (2001: 2)

> "This book is a description of the grammar of modern Standard English, providing a detailed account of the principles governing the construction of English words, phrases, clauses, and sentences."

There is nothing unique to English underlying these remarks. Parallel declarations could be found for a multitude of other languages. I conclude that linguists tend to speak rather

casually of natural language *words, phrases, clauses and sentences*. They do not feel obliged to justify the existence of these linguistic entities nor to give an account of their ontology before proceeding to their descriptive or theoretical business.

So essentially no linguist seriously doubts things like words and sentences exist, although Noam Chomsky did in effect once deny it. Postal (2004; chapter 11: 296-305) provides an uncomplimentary analysis of this preposterous, idiosyncratic denial. Given that the linguistic entities at issue exist, my simple-minded view is that the ontological question about natural language reduces largely to the issue of the ontological nature of such entities. I focus on sentences for simplicity. So what kinds of things are they? Two distinct classes of answer have existed. The sociologically overwhelmingly dominant one at every period, call it the *naturalistic* view, claims that sentences are an aspect of the physical universe. The other, called *platonist* or *realist*, associated in more recent times particularly with the work of the American philosopher Jerrold J. Katz, takes sentences to be abstract objects, things lacking temporal or spatial properties and not entering into causal relations.[1] It thus takes sentences to have the same nonnaturalistic ontological type as integers, sets and propositions. Various specific, technical things have been said in favor of a platonist view of natural language. One is advised to consult Katz's works cited in note 1 and, perhaps, also mine strongly influenced by his thinking.[2]

But to me, the most fundamental support for a platonist view comes from the fact that it is coherent, whereas the currently popular biolinguistic version of the naturalistic position does not come close to being such, for reasons turned to presently.

Two different versions of the naturalistic view deserve mention. Fifty or sixty years ago, many American structuralists as well as for example the British linguist Robert M. W. Dixon, claimed that sentences were just utterances produced by physical activations of the human vocal apparatus. I doubt anyone advocates an utterance-based view of sentences today, and for good reason. Specifically, each utterance has space coordinates, as well as a temporal beginning and end; sentences have none of these properties. Also, there are immensely more sentences than there ever could be utterances. So at best, utterances qualify only as tokens of a tiny subset of sentences.

While current linguistic culture does not take sentences to be utterances, it shares with such views an underlying drive, call it the *naturalistic imperative*. This requires that specific ontological assumptions about natural language must take it to be part of the physical world. The currently dominant view takes sentences to be psychological or biological entities, aspects of the human mind or brain. This is no doubt considered by many a more sophisticated view than the utterance one. But ontologically this is true to an irrelevantly minor extent. For the mind/brain view exhibits the same flaws as the utterance ontology for sentences. Whatever goes on in mind/brains also is spatially specific, has a temporal beginning and end, and is vastly more limited in scope than the collection of sentences. So where the utterance view conflates sentences with *external* tokens, the currently popular one conflates them with *internal* tokens, albeit of a never specified nature.

What motivates the naturalistic imperative? Perhaps there lurks the fear that without it, linguistics would fail to qualify as a *science*. The idea might be that if linguistics is not a naturalistic inquiry, it cannot be intellectually serious, threatening inter alia its access to government grants; what a horror! Such a view must ignore the nature of logic and mathematics. No one could claim these are not serious domains, but few would argue that they could be, still less have been, shown to have a naturalistic basis.

For instance, were logic a matter of contingent empirical truth, what features of the physical universe could it involve empirical facts *about* and why does it have the prescriptive force that it is taken to have? One cannot actually violate physical laws; otherwise they would not be such. But people violate various principles of logic regrettably often.

What reasoning could, for example, characterize the principle of modus ponens as something of the same order as the second law of thermodynamics? If though other indubitable domains of reality are nonnaturalistic, natural language *might* also be and the field which studies it could be concerned, like logic and mathematics, with a domain of nonnaturalistic truth. Given this possibility, one might expect those most energetically advocating a naturalistic view of natural language to have created a serious literature

justifying the naturalistic view, specifically as against a platonist view. Perhaps it is just ignorance, but I am unaware of any such literature. While the vast majority of contemporary linguists accept, I suspect, some version of the naturalistic view, that acceptance seems to me to represent a largely unquestioned culturally dominant conformism.

Current instantiations of the naturalistic view, driven in particular, by the enormous influence of the views of Noam Chomsky must assume that sentences somehow represent brain features or activities. This second variant of the naturalistic view is in one sense *worse* than the utterance view. Despite its faults, the utterance view did at least provide a transparent notion of the ontology of sentences. But I defy anyone to determine what Chomsky-style biolinguistics takes to be their specific biological ontology. I have not found a clear statement in his writings.

The failure of biolinguistic views to specify the ontological nature of sentences is not mere theoretical *incompleteness*. The lack of clear sentential ontology in the work of those taking a biological view of natural language rather partly obscures the fact that the view simply makes no sense.

The reason for biolinguistic incoherence is this: Internal to the descriptive and theoretical work linguists actually *do* when not discoursing on ontological issues, essentially everything said about natural language sentences involves taking them to be abstract objects, specifically, various kinds of set-theoretical objects, e.g. syntactic trees. But sets do not occur in brains, or anywhere else in the physical universe. So questions about the physical properties of sets are nonsensical. Are sets affected by gravity? Is the square root of 9 heavier than the sum of 11 and 7?

So to say that sentences are brain objects, while describing them as set-theoretical objects represents an inherent incoherence merely *highlighted* by Chomsky's recent decade appeal to the notion *Merge,* claimed to be the core of natural language. This is explicitly defined as a set-theoretical operation.[3] Since sets occur neither in time nor in space, taking any set-theoretical object to be brain-based, the essence of Chomsky's ontological doctrine, inevitably leads to contradiction. Another form of the contradiction

arises from Chomsky's talk of Merge originating in a biological mutation, embodying the confusion of assuming a set-theoretical operation could have a direct causal relation to a spatially and temporally occurring organism mutation event. This is as senseless as saying that sunspots caused the square root of 169 to be 13.

If a biolinguist adds, as Chomsky always has, that natural languages have *infinitely* many sentences, a coherent claim if sentences are abstract objects, further contradiction is inevitable. From the view that each of the infinitely many sentences is a brain object arises the contradictory entailment that a finite brain can incorporate infinitely many objects or actions. A brain would have to be a sort of Hilbert's Hotel for sentences. But Hilbert's Hotel is an infinite abstract object unconstrained by physical limitations, not a physical thing like a brain, and thus its nonfinite property yields no contradiction.[4]

Or, if, as sometimes stated, the claim is instead that the brain merely incorporates a finite biological coding which has infinitely many only *potential* outputs, contradiction is seemingly avoided. But this is intellectual sleight of hand since almost all 'potential sentences' forming any infinite set are not even *possible* biological objects. Think of a sentence which is four hundred trillion words long. The 'potential sentences' are at best abstract objects occurring nowhere in space and nowhere even in *future* time. So the sleight of hand cannot *save* the biolinguistic version of a supposed infinite natural language. Rather, it *abandons* a naturalistic position for a *surreptitious* platonist one.

Astoundingly, as pointed out in Postal (2012), Chomsky (2012: 91) *admits* the incoherence of his biolinguistic position, and essentially recognizes the set-theoretical nature of actual linguistics. But he then says that we have to accept things that don't make sense, like sets. This laugher recognizes the contradiction between claiming natural language reality is biological, while taking sentences to be set-theoretical. But even in the face of admitted grotesque contradiction, the naturalistic drive is so strong for Chomsky that he would rather accept incoherence and promulgate it to others than abandon it. And as justification, he just absurdly denigrates those elements, sets, which are the foundation of work in logic, mathematics and even his own linguistics.

Normally, the fact that a position does not make sense suffices as argument for abandoning or modifying it to eliminate the corrosive assumptions, Chomsky's a priori commitments lead to an infantile demand *that others accept his incoherence*. One might ask him what *other* people's incoherence *he* accepts and why, when Gottlob Frege received Bertrand Russell's observation that his set-theoretic ideas yielded contradictions (see https://plato.stanford.edu/entries/russell-paradox/), he didn't just reply that one needs to accept things that don't make sense, like sets.

What no doubt especially drives the view that linguistics must be concerned with something mental or biological is a profound confusion between language and knowledge of language. This confusion underlies the endless talk in current linguistics about language learning and its preconditions in human nature. One can and should grant the banality that knowledge of language is something mental/biological without accepting the dogma that therefore language itself is. Think of the analogy with logic. See the discussion in Part III.

In contrast with the incoherence of a naturalistic view of natural language, in platonist terms there is of course no contradiction between the set-theoretical nature of actual linguistics and the basic ontological view that sentences are abstract objects.

So if for some arcane reason one refuses to accept things that don't make sense and stubbornly seeks a *coherent* ontology of natural language, the bottom line is as follows. If one can't accept identifying sentences with utterances (also a view incompatible with the actual set-theoretical notions linguists appeal to), one must, as far as I can see, accept a platonist view. But see Part VII for discussion of a purported third alternative.

Part II On the Internet Nobody Knows You are a Space Alien Lizard

David Beaver (DB) in his recent post about the linguistic abilities or lack thereof of animals cited the assertion by Noam Chomsky (NC) in (1): [5, 6]

 (1) "[...] if someone could show that other animals had the basic property of human language, it would be of very little interest to the biology of language, but would be a puzzle for general biology."

DB disagrees with this claim and ends up stating:

(2) "Where I disagree with him is in the general principle he invokes, which seems to imply that even animals producing and comprehending grammatically correct English would be of no consequence for linguistics. Such a conclusion would be ludicrous."

Now, while I have rarely over the last decades defended any claim of NC's and have quarreled with many, the substance of (1) seems to me more or less potentially sound, at least up to the *but* clause, which I ignore. Moreover, one notes that in concluding as in (2), DB has a bit changed the terms of reference without warning. NC's claim was about the biology of language and DB's is about linguistics. Of course, some, including NC, largely identify these two, but such an identification has never been justified and is, I would argue, nothing but a category mistake, confusing inter alia language and knowledge of language.

Here then is how I see (1) and (2). (1) might be sound for uninteresting reasons. If one finds an animal with the same linguistic competence as humans, one with NC's point of view would be free to 'account' for that in the same way he proposes to 'account' for human competence. Namely, he could simply say that the innate linguistic organ he has posited for humans, whose nature he has never specified in any biological terms, and which has no known physical properties, is present in the relevant animal as well.[7]

If that account ultimately succeeds for humans, there is no reason why it would necessarily fail for the animal in question. Under these assumptions, the animal simply provides another subject, no different in linguistic essence from a human, and that would yield a situation which is genuinely linguistically uninteresting, whatever thrill it might

provide zoologists. Of course, if NC's innateness views are wrong, then finding a linguistically competent animal might show something, e.g. that the general skill set needed to learn a language is available more broadly than in humans.

I don't care much about (1) because I do not believe linguistics, or more importantly, its subject matter, language, have anything much to do with biology. But (2), which mentions linguistics, is another matter and I consider it deeply wrong. Two thought experiments can show why the issue of nonhuman linguistic competence is essentially irrelevant to the understanding of natural language.

First, let us briefly try to give a sense of a genuine linguistic issue, call it ISS. The point will be that ISS is such that no discovery about animal linguistic competence could bear on it in any way different than human linguistic competence does. Take ISS to be an issue which exists regardless of the specific theoretical assumptions cited to illustrate it.

Seuren (1985) cited (3):

(3) John and every woman in the village want to get married.

About this, he claimed:

(4) Seuren (1985: 22-23)

"In (10b) *every woman* cannot take scope over the whole remainder of the sentence; as a consequence it cannot mean that for every woman in the village John and that woman want to get married to each other. Its only possible reading is the one in which John as well as every woman in the village want to get married. This is explained in principle by the theory that quantifying into a co-ordinate structure is ruled out by the Co-ordinate Structure Constraint (Ross, 1967)."

So take ISS to be the issue of what principles of language determine that the universal quantifier phrase *every woman in the village* in (3) has the scope that it does. According to Seuren, these principles included Ross's (1967) coordinate structure constraint, which in Seuren's framework or that of May (1985), would preclude a required quantifier lowering into/quantifier raising out of the coordinate phrase. This claim, even if right,

leaves things mysterious however, since it is then not obvious how *every woman in the village* can have any actual scope at all. For its scope seemingly includes material external to the coordinated subject, the predicate complex *want to get married*.

We are not going to solve ISS here, since our interest is in nonhuman language possibilities and actualities.[8] Turn then to the two thought experiments. One harks back to the 1983 science fiction television series V. In this, large, evil lizard-like space aliens try to take over the Earth. They don't normally look like lizards though because of a fake outer coating which gives them human appearance. Suppose then, as a thought experiment, that someone, say Charles Manson, was a creature just like the V series space alien lizards. One then must accept that nonhumans know and can use English just like real humans (assuming there are any). And that tells us what about ISS? Evidently, nothing. Just how could Charles Manson's being a space alien lizard instead of an Earth mammal offer any insight into ISS. Nor would it matter if he had been a raccoon or a crocodile in human form.

The second thought experiment appeals to the known phrase "On the Internet nobody knows you are a dog", which comes from a New Yorker cartoon found on page 61 of the July 5, 1993 issue, available at:

http://www.jeffsandquist.com/OnTheInternetNobodyKnowsYouAreADog.aspx.

The cartoon shows a dog sitting at a computer terminal with another dog in attendance and the phrase is the caption. So simply assume the cartoon is realistic...suppose that all the postings and messages on some website, say Salon.com, have in fact been written by canines. That might have political implications, but as far as ISS is concerned, it means nothing. One learns and can learn in principle nothing more about ISS by the discovery that there are English-knowing dogs than by the discovery that there are English-knowing space alien lizards. It just doesn't matter for the nature of language.

There is one case of putative linguistic ability in nonhuman animals not covered by the thought experiments. Suppose one finds an animal who is shown somehow to know some variety of some *hitherto unknown natural language*. That would be of more interest, exactly as much as finding a hitherto unknown language known by some

humans. Maybe this could contribute to linguistics, but if so, it won't be because the language is known by a nonhuman, but because the language itself can teach us something.

However, with thousands of known languages available for study, we haven't really been able to understand that much. Why would just finding one more be likely to change things, regardless of whether that language is known by space alien lizards, canines or simply a group of people not previously contacted or known. For what linguistics lacks is not languages to study but insight into them. And that won't be provided in any clear way by further linguistic creatures, regardless of whether they are people, gorillas or lizards. Underlying these remarks is the view that language is entirely distinct from biology, just as mathematics, set theory and logic are. If we discover a crocodile with the same knowledge of mathematics as the best human mathematician, it won't inherently help determine whether Goldbach's Conjecture is true, and a set theoretical expert gerbil will not thereby provide any insight into the truth of the Continuum Hypothesis. [9,10]

Part III. Ontological Questions about Natural Language and Some Answers

A. Remark

The question/answer format of Part III was suggested by the style of the contribution to the Soviet linguistic controversy of the Georgian linguist Iosif Vissarionovich Dzhugashvili (also known as Joseph Stalin).[11]

B. The Playing Field

These remarks are concerned with several positions about the ontological character of natural languages. It has the least to say about the currently arguably dominant one, the so-called biolinguistic view advocated for many decades by Noam Chomsky. According to this view, natural language is an aspect of human nature, uncontroversially no doubt an aspect of the physical world. This position, already discussed in the two previous discourses, needs little more attention. It's incoherence shows that it is not even a real candidate for an underlying basis for linguistics. Notably, the incoherence has even been admitted at least once by Chomsky himself.[12]

A second position about natural language reality is the platonist one according to which natural languages and the sentences which compose them are abstract objects. This view, which I favor, was worked out most thoroughly by the philosopher Jerrold J. Katz; see the works cited in note 1. While not accepting Chomsky's biolinguistic view, the platonist position is criticized in several papers by Barbara C. Scholz and Geoffrey K. Pullum (e.g. 1997, 2010; Pullum 2011, 2013, 2015). They have claimed there is a third position, involving a type of object, perhaps not previously invoked in linguistics, called a *mind-dependent abstract object,* and asserted that natural languages are precisely this type of thing.[13] I suggest that on the contrary natural languages are not of this order. First, it has not, I believe, been shown that there is a coherent notion of mind-dependent abstract object. Second, even if such objects did exist, it has not been shown that there are any facts which would motivate appeal to them in accounting for natural languages. These points are defended in Part VII.

C. Questions and Answers

(1) Question: Linguists speak about sentences and languages. Do these terms denote real things and if so, what are their relations?

Answer: Yes, certainly these terms denote linguistic realities.

(2) Question: How can you support this view?

Answer: As for sentences, first, take a language, e.g., Parisian French. Imagine a possible world W in which none of the sentences *Je vous aime*, *Il pleut*, *Il est defendu de rien laver dans le bassin*, *Je peux pas le piffrer*, etc. ad infinitum exist.[14] What could 'Parisian French' denote in W? Evidently, nothing at all.

Second, it is an incontrovertible fact that linguistic studies concern themselves with sentences. So the existence of such is in linguistic practice not at all controversial. For instance, if one looks at any work on the nature of French, e.g. the insightful traditional Martinon (1926) or the goldmine of facts about French negation represented by Muller (1990), they are concerned with the properties of French sentences. One does not find any works putatively about the French language which deal not with French sentences but with e.g. French perfumes, clothing styles, crepes or wines.

(3) Question: Even if accepted, answer (2) only really concerns the existence and relevance of sentences. It does not say what natural languages are or how they relate to sentences. In particular, although this is your view, answer (2) does not argue that natural languages are collections of sentences. And, notably, one finds the following denial that natural languages are collections of expressions:

(4) a. Pullum and Scholz (2010: 5)

"There can be no doubt that one factor tempting linguists to accept infinitude is the ubiquitous presupposition that a language is appropriately given a theoretical reconstruction as a collection of expressions. This is not an ordinary common-sense idea: speakers never seem to think of their language as the collection of all those word sequences that are grammatically well-formed. The idea of taking a language as a set of properly structured formulae stems from mathematical logic. Its appearance in generative grammar and theoretical computer science comes

from that source. It is alien to the other disciplines that deal with language (anthropology, philology, sociolinguistics, and so on)."

b. Pullum and Scholz (2010: 5-6)

"One way to put this (and Chomsky comes close to saying this when he includes "intensional" in his characterization of 'I-language') is to say that the goal of a grammar is not to reconstruct a language extensionally, as a collection of containing all and only the well-formed expressions that happen to exist; rather, a grammar is about structure, which should be described intensionally, in terms of constraints representing the form that expressions share. Linguists' continued attraction to the idea that languages are infinite is at least in part an unjustified hangover from the mathematical origins of generative grammar."

So the claim that languages are collections of sentences is really in doubt is it not?

Answer: It might be if the authors of (4) had given any arguments for this rejection but they did not. They state, no doubt accurately, that the view of languages as collections of sentences is not a commonplace idea, that is, among nonlinguists. It is also not a commonplace idea that energy and matter are fundamentally the same thing or that the cardinality of the class of even natural numbers is the same as that of the class of natural numbers. But X not being a commonplace idea obviously does not entail that X is not the case. So talk of what is commonplace cannot show anything.

Further, the authors claim that the origin of the idea that languages are collections lies in mathematical logic, that it is foreign to fields distinct from linguistics which study language and that it is merely an atavistic leftover from earlier generative work. Again, these claims fail to address substantial questions about the nature of natural languages and so can in no way undermine the view that natural languages are sentence collections.

The passages grant that there are expressions, that they have structure and assert that the job of a grammar is to characterize the nature of the structure that expressions share. But this view is too vague to represent a coherent alternative to the view that languages are collections of sentences. For instance, no doubt elements from distinct languages share properties, e.g. the English/French pair in (5):

(5) a. The gorilla was pursued by a lion.

   b. Le gorille a été poursuit par un lion.

But it is not the task of any individual grammar to describe these similarities because (5a, b) are elements of distinct languages. And the relevant notion of distinctness is given precisely by the set-theoretical view of languages as collections of sentences, since one can then say that (5a, b) belong to different highly disjoint language-defining collections.

The view in (4) leads the authors to feel that it is illegitimate, unimportant or pointless to ask how many well-formed expressions there are. Without exactly rejecting the idea that languages are infinite, they strongly suggest that all size questions, hence the infinite size question, are misguided and not the concern of linguistic theory. I address this view in Part IV.

One observes that while passages like (4) deny that natural languages are collections of sentences, they fail to provide any positive alternative. Not so surprisingly then, in a work subsequent to this rejection, Pullum (2013), seemingly unwilling to advance any clear notion of what a natural language is at all, is reduced to denying that natural languages can be scientifically reconstructed. See the disussion of his remarks in the answer to question (13) below.

(6) Question: Are sentences abstract objects?

Answer: Yes. More precisely, they are sets, a view supported by almost any work on syntax, where sentences are treated as trees, graphs, etc., set-theoretical objects par excellence. This point needs defense against phonetic facts; see Part VI for such.

(7) Question: How can one tell?

Answer: As noted earlier, the essential properties of abstract objects are that they are not located in time or space and cannot enter into causal relations, for instance, cannot destroy anything or be destroyed. Thereby they are independent of the laws governing the physical universe. Investigation of particular natural language sentences, e.g. (8), shows that they manifest these properties.

(8) The number of erroneous views about that is quite extensive.

No one even tries to find the spatial or temporal coordinates of any given sentence. How could one? There is no way this sentence threatens anyone or anything because it enters into no causal relations. So if someone tries to assign you the task of destroying it, it would be wise to reject that challenge and similar challenges to do the impossible. These properties of sentences contrast massively with the properties of physical objects.

(9) Question: The answers to questions (1), (3), (6), and (7) together mean that natural languages as wholes are abstract objects? That is, collections/classes/sets and in general set-theoretical objects are abstract objects par excellence. But don't grammars of natural languages usually talk about the pronunciations of sentences? And don't pronunciations inherently denote movements of the human vocal apparatus and the physical things (sounds) such movements produce?

Answer: Yes, certainly grammars commonly do talk about pronunciations. For instance, Riegel, Pellat and Rioul (1994: chapter II) devote more than twenty pages of their grammar to French phonetics.

(10) Question: But displacements of vocal apparatus and physical sounds are found in time and space, are caused and can cause. Therefore they are not abstract objects are they?

Answer: No, they certainly are not.

(11) Question: So aren't the sequence of answers so far including those to (9) and (10) fundamentally contradictory?

Answer: Not necessarily.

(12) Question: How can that be?

Answer: Part V is devoted to answering this question and to a discussion of the relation between phonetic tokening systems and natural language sentences.

(13) Pullum (2013: 504) states:

> "It seems to me that the notion of 'a language' should not be regarded as scientifically reconstructable at all. We can say in very broad terms that a human language is a characteristic way of structuring expressions shared by a speech community; but that is extremely vague, and has to remain so. The vagueness is

ineliminable, and unproblematic. Human languages are no more scientifically definable than human cultures, ethnic groups, or cities. The most we can say about what it means to say of a person that they speak Japanese is that the person knows, at least to some approximation, how to structure linguistic expressions in the Japanese way (with object before verb, and postpositions, and so on). But in scientific terms there is no such object as 'Japanese'."

Question: Do you agree with this rather nihilistic account?

Answer: No, I find it mostly quite unacceptable.

(14) Question: Why?

Answer: Most strikingly, there is no argument for any of its complex claims. They are simply asserted. And this is not due to a failure of mine to quote more of the relevant surrounding discourse. A justified response could thus without argument simply negate each claim. But one can do better than that pure debating approach.

First, turn to the claim that there is no such object as Japanese. Since neither Pullum (as far as I know) nor I are students of Japanese, let us, while preserving the logic of the discussion, replace *Japanese* by *English*. There is a sense then in which I agree with a claim to the effect that 'there is no such object as English'. But the reason why that claim is correct has ultimately no ontological significance. Basically, it is true only because there are vastly more natural languages than there are *names for natural languages* and the use of terms like *English, French, Arabic*, etc. is a necessarily crude attempt to name untold numbers of distinct languages which, however, *intersect massively* in their collections of sentences.

For example, it is perfectly correct *in ordinary terms* to say that Pullum and I both speak a language called English. That is correct because there are endlessly many sentences which are elements of the collections forming both of our individual languages.[15] But there are also many which are (I gather) part of his British English and not part of my American English and no doubt vice versa. Examples of the former would be:

(15) a. This will be a negotiation and I am sure the Spanish, Portuguese or Greek tourist industry will want to resist any actions that will dissuade anyone traveling to their country on holiday.

http://www.telegraph.co.uk/news/2016/09/11/british-holidaymakers-could-be-forced-to-pay-to-travel-to-europe.s

  b. The committee are now hearing from young people's representatives on how to combat negative body image #youthselect

https://twitter.com/YourUKParl

  c. 'It's a cheap shot!' Diane Abbott lambasted live on BBC for rubbishing May's Brexit speech.

http://www.express.co.uk/news/uk/857766/brexit-news-diane-abbott-bbc-latest-may-speech-eu-theresa-may-delay-brussels-europe

  d. Boris's call to 'let the British lion roar' gets the conference hall to its feet as a volley of speeches banishes Brexit glumbucketry, writes QUENTIN LETTS

http://www.dailymail.co.uk/home/index.html

Example (15a) is ill-formed for me because the construction with the verb *dissuade* and following DP + gerund requires a *from* before the gerund. Example (15b) is ill-formed for me because nominals formed on the noun *committee* require singular agreement. Example (15c) is ungrammatical because I have no verb *rubbish*. Finally, (15d) is not part of the language I know, which contains no word *glumbucketry*.

These and multitudes of other cases show there is no single language called English which covers both British and American English viewed as collections of sentences, since collections are extensional, defined uniquely by their membership.

And the mismatch between language names and individual languages is far more massive than indicated by examples like those just given. It is quite clear that even within a single language community the number of distinct languages covered by a term like *English* is extraordinary. In fact, new expressions and new associations of meanings with previous expressions constantly appear, e.g. those like *Brexit*, *woke*, etc.[16] Each such expression defines a distinct collection of sentences and hence a language distinct from a

collection not containing it. And of course over time expressions vanish. For instance, there used to be a hostile term of male address in American English *bub*. I can't imagine anyone under 95 using it today.

Moreover, even within what is called contemporary standard American English, there are real differences. I spent most of my life in the greater New York area but now live in a suburb of Seattle in the state of Washington. I was surprised to note that interstate highways, whose names involve numerals, e.g. 87, 284, and other roads whose names involve numerals are systematically pronounced here with the definite article. So the major north/south interstate highway here is *the 5*. But no one in New York in my experience would ever have referred to the major north/south interstate highway there as *the 87*, it was simply *87*.

So for this ontologically uninteresting reason, Pullum's final claim about Japanese is correct. But it tells us nothing of importance about the ontological status of natural languages. More precisely, an ontological argument can only be formulated by focusing on individual languages. There is no theoretical reason why linguistics should try to characterize some unique object to which, for historical and *partial* intelligibility reasons, the name English is currently applied. Practically of course, when one applies to be an English teacher in some country, it is sensible to expect a demand to show some competency in languages currently denoted by the term *English*.

As for the rest of claims in (13), I find no reason to accept any of them. For instance, it is stated flatly that natural languages are no more scientifically definable than human ethnic groups, cultures or cities. But no hint of support is given that the latter are not scientifically definable, still less any that the task of defining natural languages is equally impossible. For instance, it might be true that the latter three have the asserted property but not true that natural languages do. Most of (13) strikes me as a curious cry of undefended theoretical pessimism.

(16) Question: Languages contain sentences which contain morphemes, which can be regarded as containing sequences of phonemes. These sequences involve some number of phonemes. Abstractly, there is no reason for there to be a bound on phoneme sequence

length. But is it not a fact that all known morphemes in all known languages are rather short (possibly none is longer than twenty phonemes and surely none is longer than fifty phonemes)?

Answer: Yes.

(17) Question: Shouldn't linguistics provide an explanation of the length limit on morphemes?

Answer: Yes, linguistics in the broadest sense should explicate this class of facts.

(18) Question: Do you claim there is a platonic linguistic law of something like the form:

(19) For integer n, no morpheme is such that its phonemic sequence contains more than n phonemes.

Answer: No.

(20) Question: Why not?

Answer: When one speaks of the phoneme length of attested morphemes, one is, obviously, talking about elements of languages *which have been learned*. This reveals the need to make the following distinctions:

(21) Linguistic Categories

  a. Natural Languages

  b. The subcollection of a.: Learnable (by humans) natural languages

  c. (i) The subcollection of b.: Already learned learnable natural languages;

   (ii) The subcollection of b.: As yet unlearned learnable natural languages

  d. The subcollection of a. disjoint from b.: Unlearnable natural languages

While the notion of an *unlearnable* natural language will no doubt seem bizarre or even incoherent to many, it is neither and it is not hard to give a technical argument for the existence of such systems; see the discussion of unlearnable natural languages in Langendoen and Postal (1984: especially chapter 6), in Postal (2003) and the remarks in Part IV after (8).

  Just as class (21b) is properly included in class (21a), the classes in (21ci,ii) are both properly included in class (21b).

Platonic linguistics claims that the class (21a) denotes something which does not exist in time or space and which is hence unlinked to human nature and history. But the subsetting principle yielding (21b) is a function of natural world human nature and the characterization of classes (20ci,ii) must invoke both human nature and human history. In 50BC when Caesar crossed the Rubicon, every variety of English was a learnable language but none was yet a learned one.

So any actual spoken language L of course represents something determined by the principles defining natural languages. But the possibility of focusing on L in particular in this world depends on the principles defining learnable (by humans) natural languages.

From this point of view, it is far more plausible to attribute the phoneme sequence length limit on attested morphemes to (21b,ci) than to (21a). For instance, there is no possibility of anyone learning a morpheme whose phoneme sequence has two hundred elements. Thus it would be redundant and a violation of Occam's Razor to propose a principle like (19) to explain it. While I cannot specify what the actual learning features are which drive attested phoneme length to be very small indeed, it seems entirely plausible that there are such.

(22) Question: For centuries, linguists have studied a phenomenon called 'language change'. Is that phenomenon real?

Answer: Yes, of course.

(23) Question: But since languages are, according to you, abstract objects, and those things supposedly do not exist in space or time, don't and can't change, doesn't the answer to (22) contradict your view?

Answer: No.

(24) Question: How not?

Answer: Language change is a real phenomenon but does not involve changes of languages in the sense of collections of sentences. Except when speaking metaphorically, one set of elements cannot be taken to change into another. Rather, in accord with the distinctions in (21), so-called language change is change in people and groups of people.

That is, certain people P and/or their children come at time T to have knowledge of a (slightly) different language system than the one P had knowledge of earlier.[17] In the area of vocabulary, as alluded to above with mention of *Brexit*, this happens with vast frequency…think of words like *iphone, router*, *dox*, etc. unknown a few years ago. In technical fields, government, etc., introduction of new words is often nearly continuous. But it is just not appropriate terminology to think of such matters as involving changes in languages. Historical change involves an evolution of social recognition of one system of linguistic knowledge, involving language X, to recognition of another knowledge system of a distinct (to some degree) language Y. But language X is unchanged. Thus the study of language change inherently must take into account both the nature of language and the nature of linguistic knowledge as well as the properties of the creatures who can develop such knowledge. But language change has, as far as I can see, no known implications for the nature of language as such.

Part IV Sentence Size and Language Size

A. The Infinity Controversy

In passages like those in (4a, b) of Part III above, Pullum and Scholz have rejected the idea that natural languages are properly reconstructed as collections of sentences. They offer no clear alternative but invoke the idea that the goal of grammars is the description of the structure of sentences. They assume this idea is independent of the view that languages are sentence collections. Thereby, they appear to imply that linguists need not concern themselves with questions of the number of sentences in languages, specifically, no attention to whether that number is some infinity.

I believe though that this position ignores strong interrelations between sentence structures and the size of sentence collections because any conclusions about the sentence structures of a particular language depend in part on being able to document sentences of certain lengths. That holds because some structures simply do not exist in sentences of less than length *N* for some *N*.

For instance, it is no doubt universally believed that some English sentences contain conditional clauses. But the truth of that belief requires some sentences to have at least four words, because I think that (1) is as short as a conditional sentence can be:

(1) If so, then what.

So to claim English has no sentences greater in length than three words would claim English has no conditional clauses.

Similarly, a conclusion that the English *Negative Inversion* construction illustrated in (2a) subsumes cases where the phrase fronted before the auxiliary is a *Not only* expression is only possible if there are sentences of length seven, as I believe (2b) is as short as such a Negative Inversion case can be:

(2) a. Never will she sing.

   b. Not only did Sandra attend, she spoke.

So to deny that English has sentences longer than six words is to deny that Negative Inversion construction permits *Not only* phrases to be the extracted expression.

Consider too the following query. Can the negative polarity item *in years* occur inside a superlative nominative expression? Reaching a positive conclusion requires consideration of sentences with at least nine words, since, as far as I can see, nothing shorter than (3) will be grammatical.

(3) The nastiest pundit who has resigned in years agrees.

These examples and as many more as one would want to bother to point to show that sentence structure cannot actually be regarded as independent of sentence complexity/length, for the simple reason that the very existence of specific constructions is dependent on the existence of sentences of certain lengths. Explicitly or implicitly, to answer questions about conditional clauses, Negative Inversion, the distribution of the NPI *in years*, etc., one needs to answer questions about sentence complexity/length. And questions about sentence length define in effect questions about the cardinality of collections of sentences.

The point is still stronger. English has coordinated clauses, both with *and* and with *or*. But the scope of such coordination is not given a priori. One should ask, for instance, whether examples like (2b) can be coordinated with *and*. To answer positively one needs to document sentences like (4):

(4) Not only did Sandra attend, she spoke and not only did Vic attend, he responded.

Recognizing the grammaticality of (4) requires accepting some sentences of at least length fifteen. Further, consider whether sentences like (4) can be embedded (with the needed complementizers) as complements of *it is doubtful*. The answer is positive, but to show it requires accepting sentences of at least length twenty:

(5) It is doubtful that not only did Sandra attend, she spoke and that not only did Vic attend, he responded.

Next ask whether sentences like (5) can be embedded as complements of the verb *assume*:

(6) Steve assumed that it was doubtful that not only did Sandra attend, she spoke and that not only did Vic attend, he responded.

Accepting a positive answer requires accepting that there are sentences of at least length twenty three.

This kind of argument straightforwardly generalizes to the conclusion that investigation of grammatical possiblities requires consideration of sentences of increasingly greater length. I see no rational way to limit grammatical questions of the sort just gone over and thus see no way to bound the length of sentences qualifying as answers to them. In other terms, the size of sentences that must be considered in linguistics is not subject to the whims of researchers and we cannot determine a priori what size sentences must be considered to answer grammatical questions.

The conclusion just reached would seem to conflict with the following view:

(7) Pullum (2013: 502)

> "Furthermore, nothing requires us to pick one of the definable sets of structures and equate it with **the human language we seek to describe**. Given a class M of eligible structures to be models for a set of constraints Γ, there will be a unique set containing all and only the structures in M that satisfy Γ; but which class we choose as our M is quite a different matter from determining the formulation of the constraints in Γ. The work of the grammarian is the latter, not the former. Stipulating a model class will be a separate matter determined **by the theoretical goal at hand.** One might want to set M to be a finite class large enough to include structures for all of the sentences that have ever appeared in The New Yorker, so that it would be possible to use the arithmetic of finite sets on the entire class in some statistical corpus investigation. For a different purpose one might want to set M to be the set of all finite trees, or perhaps even all trees whether finite or not. It would depend on the purpose at hand, and a lot of the time it would not even matter. A typical MTS grammar would therefore not say anything about how big expressions can be." (emphases mine: PMP)

I find these remarks strange. The first emphasized portion says the goal is to describe a human language. But the second and what follows seems to set that aside and allow various other goals, which could conflict with the goal of characterizing the language by

specifying some arbitrary (from the point of view of the whole language) limit on sentence (hence language) size. Overall, the view in (7) seems to come close to confusing the goal of characterizing a natural language with linguistic engineering applications since 'Some statistical corpus investigation' has as such nothing whatever to do with characterizing a language.

Further, I also find the following remark mistaken:

(8) Pullum (2015: 18)

> "What we really want, though, is a way of representing languages as having syntactic structure without committing ourselves to any motiveless upper or lower bound on the numerosity of sentences."

As has been shown, by recognizing the existence of certain structures and by accepting that certain subsets of other constructions exist, one thereby does commit to the position that there are sentences of lengths greater than various N. Each such case motivates that the lowest bound is at least N.

No doubt for any attested language L, there will have to exist some bound on sentence length N such that every construction in L is instantiated in sentences of length N or less. Such a fact would though be attributable to the constraints on language learning rather than to the nature of languages as such.

But if sentence size questions are legitimate, then it is legitimate to ask whether there is any bound on sentence complexity/length at all and thus to ask whether the sentences of a language are finite in number or not. I do not see how such questions can make sense if languages are not collections of sentences. And if languages are collections, it is certainly sensible to ask whether they are infinite collections, and if so, what size infinity.

An essentially contrary view is formulated as follows:

(9) Pullum (2015: 18)

> "The idea that languages could be antecedently discovered to be infinite rather than theoretically defined to be is of course untenable. No amount of evidence from any language could support or refute the claim of infinitude." "Human

languages have not been found empirically to be infinite. Rather, they are stipulated to contain a denumerable infinity of finite-length sentences by the facts of how generative grammars are defined. The finite-length limitation and the denumerability condition could be revised either upward or downward given appropriate theoretical alterations."

I turn to such claims in the next section.

B**. Arguments for Infinity, Bad and Good**

A major reason for the sort of skepticism about natural language infinitude seen in (9) no doubt lies Pullum and Scholz's central conclusion about what they have variously called the *master argument* (Pullum and Scholz, 2005) or the *standard argument* (Pullum and Scholz, 2010) for the infinitude of natural languages. This appeals to the unbounded extensibility of various constructions. Restating, extending and making more precise the discussion in Langendoen and Postal (1984: 30-35), they conclude rightly that this form of argument is not sound despite its repeated statement over long periods by many authors including, alas, the present one. However, the unsoundness of that argument and hence the failure of one basis for viewing natural languages as infinite is of course as such not an argument that natural languages are not infinite.

But the idea in (9) that the view that languages could be discovered to be infinite is 'of course untenable' and the claim that no amount of evidence could support or refute claims of infinitude are just asserted without argument and are not really supported by the justified criticisms of the standard/master argument. One could, specifically, argue rationally that languages are infinite by finding grounds for claiming they are not finite, which would, for example, be supported by arguing that there is no basis for concluding there is any specific length bound on sentences. Moreover, such arguments exist in the literature, and without referring to them, Pullum more or less reiterates and grants their assumptions.

Specifically, in effect restating the logic of an argument in Katz (1966:122), elaborated in Langendoen and Postal (1984: 35 ff), Postal (2003) offers an argument for

the infinitude of specific natural languages entirely distinct from the standard argument and independent of any stipulation that natural language grammars are generative.

It can be sketched as follows. For any language L, logic determines that either there is a class of longest sentences of L, or not. If not, then L is infinite. Thus one who rejects the view that L is infinite is committed to the existence of some finite bound on sentence length, call it N. But at least in languages with various sorts of repetitions of phrases, the issue of the existence of a bound N on the maximum number of repetitions arises. Postal (2003) illustrated the relevant English construction with examples like:

(10) My father's (father's (father's (father's... father died.

If English is finite, there is some bound N which gives the maximum number of possessive *father's* combinable in sequence as in (10). Since each longer sentence of the form in (10) expresses a proposition distinct from that expressed by any of those with shorter sequences (all in fact true), the existence of a bound N would mean English is unable to express an unspecifiably large number of true propositions about human mortality.

One could justify that if there is a bound N at least as great as the length of *father's father's* sequences native speakers can directly judge to be well-formed. But beyond that, there seems no nonarbitrary way to specify such a specific N. What evidence could show that the bound is N and not N+1 or N-1. It is not accidental, I think, that no attempt to characterize the needed N is known.[18]

Pullum's statement that no amount of evidence could support the view that a language is infinite thus seems to sidestep two key points. First, if there is no actual maximum length N, English is infinite. Second, an evidentiary chasm would have to be crossed to justify that the longest unattested grammatical example of type (10) is of some length N and not N+1 or N-1. The skeptic of the infinitude of languages like English owes, but Pullum and Scholz have not provided, grounds for asserting that the argument does not disconfirm the claim that English is finite.

Bizarrely, not only does Pullum not attack the implications of the argument type just gone over, he basically rehearses its abstract structure in the following remarks, which. however, do not reach the same conclusion:

(11) Pullum (2013: 503)

"This gives us a different view of a strange dispute that has come to the fore in recent years concerning whether infinitely many expressions are grammatical in human languages. It has become standard to argue from lengthenability: because a clause like *That was very nice* can be lengthened to make *That was very very nice*, or *That was very very very nice*, and so on, therefore the set of all expressions must be infinite. Otherwise we have to assert the existence of a number *m* such that although *very^m nice* is grammatical, *very^{m+1} nice* is not, and that seems absurd."

But if, as Pullum asserts, positing *any* particular finite bound on the *very*^{m} construction is *absurd*, then there is no such bound. And likewise for the *father's father's* construction. In short, he has granted in (11) that English is not finite.

Rather than accepting the reasoning just gone over, Pullum claims instead the following, where his *MTS* stands for model-theoretic grammars.

(12) Pullum (2013: 503)

"MTS opens up a third way, discussed in more detail by Pullum and Scholz (2010): the grammar entails neither consequence. The model class can be chosen

for particular theoretical goals, and can be set large enough to allow *very^{m+1} nice* or not, with no consequences for the content of the grammar."

The idea here is that a proper model-theoretic grammar will say nothing about sentence length, and that linguists can, as in the remarks in (7) above, choose a model class for the the grammar based on particular theoretical goals. But since statement (11) in effect grants that English is not finite, to go on and talk about 'third ways' which can avoid

taking any position on the matter approaches incoherence. It is rather like saying plutonium is toxic but there is a third way which takes no position on its toxicity.

What I think has gone wrong here is that Pullum has concluded that adopting nongenerative (nonproof-theoretic) model-theoretical grammars permits linguists and their proposed grammars to remain noncommital on issues of infinitude (whereas generative grammars having recursive rules determine that languages are of the same infinite order as the integers). But whatever grammars say or fail to say, if there are sentences, then their totality has some cardinality. [19]

Simply put, there is no third way because logic does not permit there to be a third way. Even if one is unwilling to take a position on sentence finiteness and language finiteness and even if one adopts a theory/model-theoretic grammar of some domain which says nothing about sentence size or language size, it remains a fact that there is either a class of sentences of some greatest size K or there is not and therefore remains a fact that there are either just finitely many sentences or not.

Pullum is thus mistaken to think that appeal to model-theoretic grammars and their models offers some ground for concluding that the determination of the specific sizes of natural languages is not a real linguistic issue. [20] And he has not discussed the relation between sentence (hence language) size and the existence of various constructions, as gone over at the beginning of Part IV.

That there is some fundamental confusion in remarks like (11) and (12) is highlighted and further supported by the fact that Pullum and Scholz have independently granted multiple times that some natural languages, specifically, English are nonfinite.

(13)   a. Pullum and Scholz (2001: 33)

"Indeed, as noted by Harris (1968: 11) spoken expressions can even incorporate random vocal noises, (of which there is arguably a nondenumerable infinity). Familiar utterances such as (15) suggest that the variety of material which can fill a categorized slot in a linguistic expression is in a sense not even bounded by the combinatorial possibilities of some finite phonetic vocabulary."

b. Pullum (2013: 508)

> "Postal notes that the items that can appear in sentences may not even be denumerable: as Zellig Harris (1968, p. 11) pointed out, in utterances like *He went ___* the complement of *go* can be any of an in-principle nondenumerable infinitude of utterable sounds."

   c. Pullum (2015: 15)

> There appears here the exact wording of (13b).

So, despite claiming in (9) that 'no amount of evidence' could bear on the infinity question, he essentially endorses the position argued at length in Postal (2004: chapter 6). That claimed that for direct discourse sentences like e.g. *Veronica screamed vtsotumskapso* the class of verbal objects is not only not constrained to be finite, it is not even constrained to be of the smallest infinite size. But one cannot coherently claim that no evidence bears on the infinity of natural languages while also granting that there are constructions sanctioning nondenumerably infinitely many components. The implications of cases like those mentioned in (13) are discussed further in Part VI. Incidentally, I have ignored here the fact that (13), especially (13a), seem to me as written to conflate sentences with their utterance tokens.

An argument somewhat similar to that based on (10) is discussed in the following:

(14) Pullum and Scholz (2003: 70):

> "Again, take the rhetorical question asked by Atkinson, Kilby and Roca (1982, 35-36): 'whether anyone would seriously suggest that there is a number, n, such that *n is a number* is a sentence of English and *n+1 is a number* is not a sentence of English.' On the basis of an assumed negative reply to the question, they take it as conclusively demonstrated that the number of English sentences is infinite. This is no conclusive demonstration. Would anyone seriously suggest that there is a number n such that a pine tree can have n pine needles but cannot have n + 1 ? The invited negative answer does not license us to conclude that there are pine trees with infinitely many needles."

Pullum and Scholz's reply to the Atkinson, Kilby and Roca argument is partly sound. Certainly, the latter's remarks did not form a conclusive demonstration or, indeed, any

demonstration at all. At best, the argument gave support to the conclusion as the most reasonable and plausible conclusion from the facts. In this, it was parallel to the argument from (10).

However, Pullum and Scholz's dismissal of the argument on the basis of the pine tree comparison is too clever by far. It ignores the highly relevant distinct overall contexts in which sentences and pine trees can be considered. Since the latter are physical objects, they are subject to the laws of nature. While no doubt an exaggeration to say that there is a specifiable number N such that pine trees can have N needles but not N+1, it is quite true that there are many numbers of needles which no existant pine tree can ever have. But in discussing collections of abstract objects like integers or natural language sentences, there is no relevant analog to the physical laws limiting pine needle numbers. If there were, claims about the arbitrary nature of any N needed to bound sentence length would collapse and arguments for infinity based on such arbitrariness would collapse with them. So in the end, because of the ontological contrast between pine trees and sentences, the pine needle/sentence analogy is a nonsequitur.

Just before their remarks in (14), Pullum and Scholz discuss a distinct sentence analogy, this time with worms, as follows:

(15) Pullum and Scholz (2003: 70)

"Or consider Stabler's reasoning in (89): 'there seems to be no longest sentence, and consequently no maximally complex linguistic structure, and we can conclude that human languages are infinite' (Stabler 1999:321). It is just as true that there seems to be no longest worm. There was a bootlace ribbon worm (species Lineus longissimus) washed ashore in a severe storm at St. Andrews, Scotland, in 1864 that proved to be 180 feet long, i.e. around 5,500 cm. But for the storm, it could easily have grown another centimeter to reach 5,501. Surely a worm n centimeters in length could always grow one more centimeter. But we cannot conclude from the fact that there seems to be no longest worm that we can conclude the population of worms on earth is infinite. It is not, because ultimately there are physical resource bounds. Worms do not constitute an appropriate

domain for the use of any principle like the Principle of Mathematical Induction."

First, it is not a fact that there seems to be no longest worm...all they indicate is that a worm can be much much longer than one might have supposed. But they themselves note the disanalogy between sentences and worms and indicate that resource bounds will guarantee there are some limits beyond which there can be no worms. Amazingly though, they fail to draw the then clear conclusion that there being no analog of resource limitations for sentences/integers (see note 18), the analogy with worms like that with pine trees can show nothing.

My conclusion is thus that, stripped of its excessive 'demonstration' rhetoric, Atkinson, Kilby and Roca's (1982) argument does support the infinitude of English.

C. The Coordinate Closure Argument

A different discussion in the literature claims that natural languages represent nonfinite collections. Langendoen and Postal (1984, 1985) argue that a language having coordination of clauses sanctions the existence of clauses of every length, finite and nonfinite. I will not attempt to sketch the argument in detail. Essentially, it claims that English in particular manifests the following property, called *Closure under Coordinate Compounding of Sentences* in Langendoen and Postal (1984: 53). Call it CCC. Informally, CCC says that for any subset K of English sentences of, in particular, the declarative type containing two or more elements there exist one or more grammatical coordinate sentences not in K whose conjuncts are the elements of K, prefixed in most cases by *and*. Precise statement of CCC presupposes that a coordination of declarative sentences is itself a declarative sentence.

So for example according to CCC, sentence set (14a) yields a number of sentences including those in (16b, c, d):

(16) a. {Strong winds are dangerous, Rabbits loathe licorice, She may be guilty}

    b. Strong winds are dangerous and rabbits loathe licorice and she may be guilty.

c. Rabbits loathe licorice and strong winds are dangerous and she may be guilty.

d. She may be guilty and strong winds are dangerous and rabbits loathe licorice.

Since the resulting coordinate sentences, here (16b, c, d), are not in K, they form other subsets of sentences governed by CCC, which then form further coordinate sentences in turn forming possible inputs to CCC, etc.

The key to the argument then is that the ever larger sets sentences including the coordinate sentences formed under CCC can be shown to correspond precisely to increasingly large subsets of the overall language. In particular, sentences formed under CCC yield sets which are themselves inputs to CCC. Langendoen and Postal (1984, 1985) then appealed to Cantor's Theorem (see https://en.wikipedia.org/wiki/Cantor%27s_theorem), which shows that for every set T, the cardinality of the power set of T (the set of all subsets of T) is greater than that of T. Given that, it is shown that English has sentences of every finite and nonfinite length and that there is no end of them.[21]

Pullum and Scholz (1997) and Pullum (2013) reject the argument in Langendoen and Postal (1984, 1985). The earlier rejection is as follows:

(17) Pullum and Scholz (1997: 31-32)

"Only one argument of this sort has been offered: the argument of Langendoen and Postal (1984) from the principle of closure under coordinate compounding. But the argument simply will not do the job. In brief, the idea is this. The principle of closure under coordinate compounding is claimed to be a universal of language. Assume for the sake of argument that it is, and that it can be stated (informally) as follows:

(3) For every collection $C = \{S_1, S_2, ...\}$ of sentences in a natural language L there is a coordinate sentence of the form 'S$_1$ and S$_2$ and ... ' in L, i.e. a sentence in which each member of C occurs as a conjunct.

Langendoen and Postal claim it follows that sentences can be of infinite length (and there are transfinitely many sentences, and grammars cannot be equivalent to Turing machines), because of a single fact: *C* could be an infinite collection. To stipulate that *C* must be finite, by adding the word 'finite' before the word 'collection' in (3), would be empirically unmotivated; no data could possibly provide evidence for it. The trouble with this argument is simple enough to see: a constructivist could adopt strict finitism (some do, though many do not), and under that position there are only finite collections. In that case the stipulation vanishes. Collections can only be finite and thus sentences can only be finite. Whether or not strict finitism is untenable (see Wright 1993 for an extended reflection on the question) is not the point here. The point is merely that what turns out to be the simplest statement of the principle of closure under coordinate compounding depends on what is assumed about the existence of the completed infinite sets of classical mathematics. Langendoen and Postal's argument only holds, question-beggingly, for those who have already rejected the strict finitist version of constructivism. Langendoen and Postal are not entitled to shortcut the debate by stipulating that only supporters of classical Cantorian set theory can play."

So the criticism is that Langendoen and Postal's (1984, 1985) argument assumes the existence of nonfinite sets. Pullum and Scholz take that to be question-begging because there are those, so-called *finitists*, no doubt in the 21st century only a tiny number (see e.g. https://en.wikipedia.org/wiki/Finitism) who deny such existence or at least consider it controversial. By not addressing a finitist rejection of infinite collections, Langendon and Postal's argument is thus supposedly fatally flawed.

Is this a fair and justified criticism of the argument we gave? It could only have been such if certain conditions were met. Most critically, the criticism could not be used

arbitrarily to attack one particular grammatical argument. It would have to be an instance of a *general* methodology M. This would have to assert (and justify) the claim that, in the absence of a confrontation within framework A of the finitist issue, any appeal in A to nonfinite collections can be rejected as unsound question begging (not only in linguistics but in any field).

Pullum and Scholz (1997) neither formulated any M nor referred to any literature which provided a characterization of such. Specifically, despite their great formal sophistication, they failed to observe the longstanding overall intellectual context relevant to finitism, which is very nicely expressed in (18): [22]

(18) Potter (2004: 70)

> "Modern mathematical practice makes use of infinite collections everywhere, at least if we take its surface grammar at face value; not only does the construction of proxies for the standard objects of mathematics in set theory that we shall outline in part II need there to be at least one infinite set in order for it to get started, but even independently of that construction mathematicians quite routinely use infinite sets in reasoning about these standard objects."

In other words, in current formal work, finitism is an extremely marginal position, to the extent it is an active position at all. This indicates that if one is going to accuse pieces of research of question-begging because of the invocation of nonfinite collections, it would not be a rational choice to pick Langendoen and Postal (1984) as the thousandth target to begin with, still less as the first.

The question-begging criticism, if appealed to at all, should arguably have initially focused on Cantor and the innumerable renowned and respected logicians, mathematicians and set-theorists who subsequently have appealed to nonfinite collections, for example, the following:

(19) Quine (1953: 93-94)

> "This axiom asserts that there is a class with infinitely many members. But in the present system such a class is forthcoming without the help of the axiom, namely, the class V, or $\hat{x}$ ($x = x$). The existence of V is provided by R3'; and so is the

existence of infinitely many members of V, namely, Λ, {Λ}, {{Λ}}, {{{Λ}}}, and so on."

(20) Stoll (1963: 85)

"We turn our attention next to the nonfinite cardinals. A nonfinite cardinal is an **infinite** or **transfinite cardinal**. If the cardinal number of a set is infinite the set is called **infinite**."

(21) Boolos and Jeffrey (1989: 3)

"A list that enumerates a set may be finite or unending. An infinite set which is enumerable is said to be *enumerably infinite* or *denumerable*. Let us get clear about what things count as infinite lists, and what things do not."

(22) Forbes (1994: 35)

"Second, in any natural language the number of words is finite, but we

will allow the lexicon of LSL to contain infinitely many sentence-letters."

None of these invocations of the infinite is associated with a discussion of or rejection of finitist claims.[23]

A just application of the standard Pullum and Scholz (1997) required of the argument of Langendoen and Postal (1984) would find all these works on formal issues to also be vitiated by question-begging. One would then expect to find writings by Pullum and Scholz where such an attack is directed at these logical and set-theoretical works or innumerable others like them. But I don't believe such criticisms exist.

Closer to home, I likewise do not believe Pullum and Scholz have ever directed their question-begging critique logic against works in *linguistics*, where it has for decades also been overwhelmingly accepted that there are nonfinite collections. See, for example, such works as McCawley (1988: 145), Partee, ter Meulen and Wall (1993: 9) and Peters and Westerståhl (2006: 72). All of these invoke nonfinite collections with no associated or referenced discussion of the finitism question. And of course Noam Chomsky and innumerable followers have invoked the infinite character of natural languages repeatedly for more than six decades, e.g.:

(23) Chomsky (2005: 11).

"An elementary fact about the language faculty is that it is a system of discrete infinity.

But I am aware of no attempt on Chomsky's part to show that finitism is incorrect.

The logic of Pullum and Scholz's (17) must take all such linguistic-internal discussions to likewise represent question begging. But I am also unaware of any attempt on their part to aim a question-begging type criticism at work in linguistics other than Langendoen and Postal (1984). Since they are frequent critics of other linguistic work, if their criticism of that work were based on general methodological principles, such a gap would be incomprehensible.

But worse, and the clearest sign that criticism (17) represents an unacceptable double standard, is that at least subsequent to Pullum and Scholz (1997), these authors did not themselves take finitism as any barrier to the reasoned invocation of the infinite. This is shown by multiple instances where they invoke infinite collections *in their own work*. This was already suggested by (13). See also:

(24) Pullum and Scholz (2006: 69)

"The Gold problem re-emerges immediately: the indefinitely large set of such derived nouns is a possibly unbounded set of strings, and under Gold's assumptions no algorithm can identify it in the limit from text."

Here the authors invoke an *unbounded set of strings*, that is, an infinite set of strings. But there is no talk of finitism.

Consider too:

(25) Pullum and Tiede (2010: 282)

"The logic $\mathcal{L}_{cp}$ has an infinite set of modalities. They are defined recursively."

And again:

(26) Pullum (2011: 3)

"In addition, remarks like "we can find various kinds of non-finite state models within English" (SS: 22–23) and the similar remark that "we arrive at subparts of English with . . . mirror image properties" (Chomsky 1956b, 1965 reprinting, p. 109) suggest a failure to appreciate that FSLs can have infinite non-finite-state

subsets. Only if such a subset can be extracted by some regularity-preserving operation like homomorphism or intersection with a regular set does it entail anything about the language as a whole."

Here Pullum without qualms writes about infinite non-finite-state subsets. There is no mention of any finitist rejection of the infinite needing to be dealt with.

And once more:

(27) Scholz, Pelletier and Pullum (2011)

"One difference between the two accounts of squaring is that the intensional one could be applied to a different domain (any domain on which the 'X' operation is defined: on the rationals rather than the integers, for example, the extension of the identically defined function is a different and larger set containing infinitely many fractions)."

Despite the absence of any mention of finitism, Pullum speaks of infinitely many fractions.

Most strikingly, Pullum (2013) invokes the infinite multiple times with no mention of finitism and no withdrawal of the finitist rejection of the argument of Langendoen and Postal (1984, 1985). Typical statements of the relevant sort include:

(28)   a. (2013: 499)

"He further proves that this yields a strict hierarchy containing infinitely many MSO-characterizable stringsets."

b. (2013: 500)

"The class of grammars Stabler defines turns out to capture what in GES terms would be the entire infinite union of the control hierarchy. That is, we get all of the stringset classes that are string yields of MSO-characterizable classes of finite n-dimensional tree-like graphs."

c. (2013: 513)

"One framework that does not entail bounded branching is that of (Gazdar et al. 1985). There the existence of a longest rule is avoided by means of a notational device equivalent to a metagrammar generating an infinite rule set."

The key here is not that these statements include the words *infinite* or *infinitely*. Various others in this work also include them, while dismissing the relevance of infinity questions to linguistics. The point here is that (28a, b, c) are statements about other work *which Pullum regards highly positively*. There is nonetheless no mention of them having begged questions about the existence of infinite collections.

The criticism in (17) turns out then to have illegitimately assumed the authors of Langendoen and Postal (1984, 1985) were under an obligation to treat logical/mathematical/philosophical questions about finitism. But such an obligation is essentially unknown or rejected in formal work dating back more than a century and in particular is unknown in linguistics, including in Pullum and Scholz's own linguistics. So their criticism (17) of the Langendoen and Postal (1984, 1985) argument invokes an unjustified standard they (quite reasonably) do not require themselves to meet. I conclude that remark (17) was just a debate style dismissal of an argument they strongly disliked but apparently had found no substantive objection to (but see the discussion below of one such putative objection).

Another unreasonable criticism of the position of Langendoen and Postal (1984, 1985) is found more recently in:

(29) Pullum (2013: 503 n1)

> "Langendoen and Postal (1984) took it to be a serious issue whether sentences of infinite length exist, and Katz (1996) concurred, in this journal; but I see no sensible issue here. The arguments given by Langendoen and Postal for the existence of infinite sentences—in fact non-denumerable sets of them—depend on unsupported assertions that natural languages, qua collections, must be closed under certain infinitary operations. No linguists have taken these arguments seriously."

Notably, this attack is quite different from that in Pullum and Scholz (1997), which criticism is not cited. Where the earlier point was that the appeal to the closure principle supposedly involved question-begging, here it is claimed that principle was unsupported.

But these newer remarks are wrong. Least importantly, Langendoen and Postal (1984) did not claim that e.g. English *must* be closed, but simply that it is. And the claim is supported by such facts as that every collection of English declarative sentences we considered does indeed map to a coordinate sentence of the claimed sort. And I still know of no set of declarative sentences which cannot be conjoined as determined by the closure principle. This fact admits of no other known account. If Pullum knows of such a set inconsistent with the claimed principle, he has failed to publicize the discovery. That would be mysterious as its display would provide a substantive counterexample to the claimed closure principle sufficient to falsify it, eliminating any motivation for the sort of nonsubstantive criticisms of the Langendoen Postal (1984) argument he and Scholz have actually offered.

Focus then on the last sentence in (29). Ignore what I consider just make believe sociology since Pullum cites no reason to believe he has polled any sample of linguists to obtain actual data supporting the claims. Worse, as shown above, he and Scholz took the argument seriously enough to provide their unjust 1997 argument against it. And still worse, read literally, this sentence entails the calumny that either the authors of Langendoen and Postal (1984) are not linguists or that they did not take their own work seriously. But worst of all, the remark is, as far as substance goes, a nonsequitur. Even if it were true, it would show nothing. Linguistics is not politics and truth is not determined by votes. Whether one linguist or every linguist took the argument seriously is irrelevant to its correctness.

In the face of the two hollow criticisms of the conclusions of Langendoen and Postal (1984), on should ask why Pullum and Scholz so strongly reject the basic argument of that work and whether they ever offer a more substantial basis for its rejection. At least part of the reason for their rejection appears to lie in their advocacy of model-theoretic grammars. The following relevant brief remarks are revelatory:

(30) Pullum and Scholz (2005: 498, n. 15)

> "The observation that model-theoretic ("nonconstructive") grammars are compatible with infinite-size grammatical expressions is stressed in Langendoen

and Postal (1984). What distinguishes our position from theirs is that they hold that natural languages are proper classes, closed under an operation of infinite coordinate compounding that renders them too large for the laws of set theory to apply. We have no space to discuss this position here, but we note one point. Langendoen and Postal claim that for every set X of sentences in a natural language L there is a coordinate sentence of L having all the members of X as its coordinates. This claim is not statable as an MTS constraint, because it is not interpretable on individual expressions. So under a strict construal of our position, Langendoen and Postal's closure generalization is not just unmotivated but actually unstatable."

Pullum and Scholz (2005) thus assume that the statements of proper model-theoretic grammars of natural languages can only quantify over the elements of sentences. The coordinate closure principle of Langendon and Postal (1984, 1985) is definitely not of this order as it references whole languages as well as sets of sentences and individual sentences. This might appear to provide a substantive ground for rejecting the coordinate closure principle.

But the argument in (30) has limited force so far *even if one accepts* their view about the constraints on individual grammars. The reason is that they failed to consider the possibility that the closure principle could be part of a universal account of the nature of language rather than a part of individual grammars. David E. Johnson (personal communication) suggests that the point can be framed as follows. Genuine linguistic universals will be stated in the metatheory defining languages and grammars, not in individual grammars.

Informally, suppose that grammatical theory defines something we can call a Natural Language Template. This would characterize the possible collections which make up natural languages and the sentences which can form them. Individual languages would then draw from the possibilities so characterized, 'choosing' some, excluding others. In such terms, the function of the grammar of an individual language is to pare

down the possibilities specified in the Natural Language Template to one particular instantiation.

The characterization of the Natural Language Template might take the form of a recursive definition, sketchily along the lines in (31), where the Closure Under Coordinate Compounding principle is given as (31b):

(31) a. L is a natural language if and only if L consists of a universe of objects called sentences and if S is a sentence of L, then S has properties $p_1...p_n$ or $q_1...q_m$.

b. If $K = \{S_1, ...S_n\}$ is a collection of sentences of L all of the same category (e.g. Declarative), then if there are structures of the form [[conjunct+S_1] [conjunct+S_2]...([conjunct+S_n])] satisfying (31a), there is a sentence in L which is the coordinate projection of K.

Here, nonessentially to simplify the presentation, I assume that every conjunct has its own conjunct marker, even if some (e.g. the first in English) are null.

Evidently, condition (31b) does not satisfy the condition which Pullum and Scholz assume for natural language model-theoretical grammars. But this is not known to be a problem since under the present speculation (31b) is part of linguistic metatheory not part of any natural language grammar. It would nonetheless impose the property defined by the condition on every language allowing clausal coordination. The latter restriction is motivated by observations that languages like Pirahã (Everett, 2005), Hixkaryána (Derbyshire, 1979) among others seem not to. See the survey discussion in Pullum and Scholz (2010: 17-20). These cases falsify the claim of Dik (1968: 1), wrongly taken as correct in Langendoen and Postal (1984: 71), that *every* language manifests such a construction.

But if natural languages are things which must satisfy the Natural Language Template, there is no obvious reason why an overall account of natural languages cannot have both the closure principle, relevant for the subset of languages like English, and the Pullum/Scholz constraint on individual model-theoretic grammars. If so, even this somewhat more substantive attack on the argument of Langendoen and Postal (1984) does not hold up.

Part V The Supposed 'Excess Language' Problem

As part of a rejection of the platonist conception of natural language, Pullum (2015) argued that the view of natural languages as abstract objects defended inter alia in parts I-III suffers from the flaw of specifying *many too many languages*, as follows:

(1) A platonist view of natural language ontology will inevitably allow as natural languages, that is, things claimed to fall into category (21a) of part III, lots of systems which are *very* unlike *attested* natural languages.

Let's call this purported flaw of platonist views the *excess language problem*. Pullum's claim (1) is, I grant, correct. But he errs, I believe, in inferring from that state of affairs to the conclusion that it represents a problem for platonist linguistics, and is, in effect, a solid part of a refutation of it. He elaborates as follows:

(2) Pullum (2015: 5):

> "One problem that as far as I can see Katz never addressed is that there are too many languages under this view, infinitely many of them being utterly unlike human languages, but all of them apparently being fit subjects for study by linguists. Katz doesn't just imply this, he insists upon it. Sets of dialects of a natural language are just classes of sets of sentences, and 'There is an infinite range of such classes, including English, French, Sanskrit, Engrench (i.e., a class of sentences with English syntactic structure but an anglicized French vocabulary), and infinitely many other languages, living, dead, unborn, conceivable, and inconceivable' (Katz 1981: 9). The set of well-formed sentences of Standard English in American spelling as exemplified in Katz's own writings is not distinguished from any of these *imaginary* (emphasis mine: PMP) languages."

Let us examine the specific examples of the excess language problem Pullum gave. The first was:

(3) Pullum (2015: 6)

> "h(English) where h is a homomorphism mapping all nouns, verbs, and adjectives to blah. (Sample text: We blah these blah to be blah, that all blah are blah blah,

that they are blah by their blah with blah blah blah, that among these are blah, blah, and the blah of blah.)"

It is no doubt true that no linguistic theory anything like one is motivated to propose would say h(English) is not a natural language. But why should it? The fact that every sentence in h(English) is endlessly ambiguous renders h(English) unlearnable. So the fact that no parallel to it is attested is nonmysterious *independent of a platonist view of linguistic theory*. Ditto the conclusion that it seems very 'unlanguage like', since that feature is partly a grasping of its unlearnability.

The second example of the excess language problem was:

(4) Pullum (2015: 6)

"The concatenation of Japanese and French (i.e., the set {wx|w is a sentence of Japanese and x is a sentence of French})."

Call this class of sentences *FrancoNipponese*. Plausibly, FrancoNipponese is also unlearnable. In any event, it is such as to place massive, useless burdens on a language learner. The latter would have to build both the grammar of Japanese and the grammar of French and then introduce a concatenation principle which I do not know is instantiated in attested natural languages. Moreover, once learned, FrancoNipponese has the distressing property that most of its sentences must say things a user would have no wish to say and might in innumerable cases wish not to say. In particular, most of the declarative sentences of FrancoNipponese represent pairs which are nonsequiturs, e.g. pairs with meanings like 'Philosophers are mammals/Most primes are rather large'. Endlessly many express contradictions. Even when the paired sentences lack these properties, they are redundant. Suppose a FrancoNipponese speaker wished to express all and only the meaning expressed by the French sentence*: Je t'adore* ('I adore you'). (S)he cannot just say that but must add on some Japanese sentence…if the latter also means 'I adore you', no harm is done, but redundancy is forced. And if the wished for expression is complex, the redundancy rapidly reaches preposterous proportions. Imagine translating (5) into FrancoJapanese.

(5) The reasons one should not be a communist are many and widely understood in some communities but not in lots of others due to mental illnesss, excess lead exposure, bad child rearing, corrupt officials and terrible educational institutions as well as lust for power, sadism and many other revolting human characteristics.

So again, it is no real mystery why nothing like FrancoJapanenese is attested and there is no documented need for linguistic theory as such to block it.

The third example of an excess language was:

(6) Pullum (2015: 6)

"The set {ii, iii, iiiii, iiiiiii, iiiiiiiiiii, . . .} containing all and only the prime-length strings of the letter i."

Call this the *i-language*. Pullum does not say what suggests to him that natural language theory would allow the i-language. But one might assume that each i-string in the set he gives could be a lexical item, and hence mapped into e.g. English or Japanese lexical items, and thus the i-language could have the grammatical structure of an attested language. This might be challenged by noting that the need to talk about prime lengths might not be feasible in a natural language. That aside though, the likelihood here is that again learnability constraints would independently account for why no analog of the i-language is spoken anywhere. First, that language presents the listener with ovwhelming ambiguities with respect to almost every sentence. Second, it suffers from the length problem for morphemes discussed in (16)-(21) of Part III above. For the i-language to have a vocabulary of twenty elements would require a morpheme of phoneme length twenty, and for a vocabulary of five hundred items one with a length of five hundred. Moreover, the burdens on memory and perception of distinguishing items represented as very long strings of *i*, e.g. of length twenty from those of length twenty one, are no doubt insuperable.

I conclude that the actual cases which Pullum cites show nicely that a platonist grammatical theory would allow many sorts of systems strikingly different from actually *attested* natural languages. But they do not in any way show that that fact is a problem for the view that natural languages are set theoretical abstract objects.

One also finds:

(7) Pullum (2015: 6)

"Linguistics is not distinguished from formal language theory or logic or theoretical computer science under Katz's view: its truths as well as its methods are entirely a priori. As he says: 'Platonists classify linguistics with the mathematical sciences, rather than with the social, biological, or physical sciences, as also about a reality outside of us and the physical world'; thus 'statements about the grammatical structure of sentences are no more empirical than statements about numbers' (Katz 1981: 23)."

Pullum finds this view unsatisfactory, as follows:

(8) Pullum (2015: 6)

"As a posteriori philosophy of science this will not do: linguists simply do not work in a way that can possibly be regarded as rendering this description true. Asserting things about the grammatical structure of sentences in arbitrary *invented* (emphasis mine: PMP) languages does not count as work in linguistics."

But claim (8) is not argued and I see no reason to believe such a grand generalization is true. It depends upon what things are asserted about the grammatical structure of what nonattested languages.[24]

For instance, to take a trivial case, Postal (2003) argued that that there is a natural language isomorphic to English which is unlearnable, just as Pullum's h(English) is isomorphic to English and unlearnable. The purpose of this argument was exactly to show that the notions natural language and learnable (by humans) natural languages are not coextensive collections. The argument ran on the statement of a precise mapping of every extant English lexical item to an image which was massively longer, so long as to be unlearnable. The point was then that there was no structural difference between extant English and the language resulting from the mapping. The only differences were in phonological lexical representations so that no known principle of language would say that the mapped to language is not a natural language. This argument might be wrong in

some way but I see nothing nondogmatic which would indicate that it *shouldn't count as work in linguistics.*

The same work also states:

(9) Pullum (2015: 6)

> "Nor does the development of true and provable theorems concerning such languages. It is certainly true that in any language where every clause contains a verb, and all verbs are octotransitive (i.e. they take eight obligatory object NPs), and subject NPs are obligatory in all clause types, and NPs minimally consist of a noun followed by a determiner, every sentence has at least 11 words; but this would not be recognized by any linguist as a syntactic discovery falling within linguistics."

But again, this is asserted without argument. Why wouldn't the conclusion just be a trivial theorem about minimal word length?

I do not see then that so far, Pullum's excess language criticism of the platonist view of natural language ontology indicates any actual problem for it. My strong suspicion is that what has led him to conclude otherwise is his failure to make the distinctions in (21) of Part III above. That is, Pullum's conclusion that the excess language problem offers a kind of refutation of a platonist view of natural languages appears to stand entirely on a failure to invoke the distinction between a language L and knowledge of L. Once this distinction is made clear, it follows that every *attested* language must satisfy not only the conditions defining natural languages but also those constraining human language learning/formation. Thus it cannot be assumed a priori that any particular features characteristic of attested languages relate to the former and not to the latter.

To summarize, the excess language problem is no threat to platonist views of natural language ontology. It merely shows that a failure to conceptually separate language from knowledge of language threatens confusion.

Part VI Natural Language Sentences and Their Tokens

Questions (9)-(12) in Part III raised the issue of the place of phonetics (and, analogously, gestures in languages of the hearing impaired) in a view of natural language ontology. This topic is addressed in the following:

(1) Pullum and Scholz (1997: 30-41)

> "Heterogenism asserts that the referents of the theoretical terms of a single generative linguistic theory are ontologically diverse. We believe that this view has some initial plausibility. Languages are structured connections between sound and meaning. At the phonetic end, no linguistic theory can adquately describe human languages without a predicate 'nasal' being involved. And statements containing the predicate 'nasal' are satisfied by concrete (and mind-independent) objects; specifically, the relevant objects here are velums (a segment is nasal if and only if the velum is lowered during its production so that the velic port to the nasal cavity is opened)."

The claim then is that natural language sentences manifest both abstract elements and concrete ones, e.g. things denoted by *nasal*. No doubt this view is initially plausible and, for that reason has probably been the universally accepted view of linguists throughout history.

That said, I think the view that there are individual things called natural language sentences which involve both abstract and phonetic properties is wrong. Under normal conditions, human articulatory movements like that described informally by Pullum and Scholz and the noises they produce are relevant because this is the standard human mechanism to assign tokens to natural language sentences. In the context of the hearing-impaired, token systems arise involving facial, hand, etc. gestures.

For the traditional and Pullum/Scholz view to be corrrect, the token forming mechanisms associated with languages must be parts of them in the same way that their syntax and semantics are. But this position is not self-evident and, notably, Pullum and Scholz (1997) do not argue directly for it. I do not know that anyone ever has. But see the discussion of their claims in (4) below. While traditionally large scale grammars

invariably attempt to give specifications about the sounds and physical gestures underlying them which characterize the pronunciations of the sentences of the language being descdribed, this tradition has as such no ontological import.

A rather different view was implicit in the discussion in Postal (1968: 6-7, 58-59), although it was partially presented in nonplatonist (mentalistic) terms which I now consider not only erroneous but incoherent. The idea focused on the notion of phonetic representation:

(2) Postal (1968: 6)

"The obvious solution to this problem, namely, the traditional realization (Sassure, 1959: 11) that phonetic representation is not a direct description of speech signal or articulation but rather part of a theory about the instructions sent from the central nervous system to the speech apparatus is, of course, intolerable to the modern linguist because it is mentalistic."

Currently, I consider the instructions to be purely abstract and not mentalistic.

Moreover, the view of Postal (1968) incorporated the same incoherence noted for Chomsky's biolinguistics in Part I above and especially in e.g. Postal (2012). This follows from a comparison of the mentalist talk in (2) with:

(3) Postal (1968: 58-59)

"Phonetic representation is properly looked upon as a two dimensional matrix which graphs phonetic segments against a set of ...*universal* phonetic features. A phonetic segment is then formally an unordered set of specifications of the values *1-n* of each of these features."

This set-theoretical account of phonetic representation is incompatible with the view in (2) that phonetic representation is mentalistic since sets are not mentalistic.

But the key current point is that confused mentalistic talk aside, (2) and (3) sharply distinguish the instructions represented by abstract phonetic representations, which should be taken to be set-theoretical parts of the abstract objects forming sentences, from the physical movements and resultant sounds such instructions bring about.

An analogy supporting the position that the physical movements of token forming systems are not actual parts of natural languages can be based on a Microsoft WORD DOC file F and its coding in a computer. File F exists entirely independently of the physical structure of computers, of the technology or technologies underlying screen displays and, most clearly, independently of printer technolgy. There is no reason why an account of the nature of Microsoft WORD files should say anything about screen technology or printer mechanisms or ink cartridges. But the normal ways of producing paper tokens of Microsoft WORD files of course involve printer technology and ink cartridges. My suggestion then is that the token mechanisms associated with the sentences of natural languages bear the same relation to them as printer technology bears to WORD files.

When the Print command is given in the appropriate context, that is, in particular where the computer is linked to a functioning printer P, F is interpreted by P as an instruction to engage in certain actions which result in ink being placed on pieces of paper. The WORD document is an abstract object which, thanks to the coding and functioning of the physical computer and linked printer, yields tokens of the file (more precisely, of the noncontrol, nonformatting parts of it). But the properties of the printer and its actions are not elements of F.

This analogy supports the view that just as printers, their technology and outputs are not parts of WORD files, articulatory phonetic and gesture mechanisms and their physical results are not part of natural languages, but rather are auxiliary token forming systems. This is not to denigrate phonetics, nor to suggest its study is unimportant nor to imply that its study is not a legitimate part of the broad field called linguistics.

Another perspective which supports the idea that the phonetic and gesture mechanisms are not strictly speaking part of natural languages derives from orthographies. Many many languages are, of course, now associated with regimented orthographical systems. These provide an alternative to the more primary token forming systems, one whose provided tokens lack the evanescence of sounds and gestures in the absence of recording and video technologies. While some people speak of spoken

languages and written languages, the ontological view urged here permits avoiding these misleading terms. What the terms *spoken* and *written* then denote is not languages but only token forming systems associated with them. If e.g. articulatory systems were really parts of natural language, it is hard to see how one could communicate in natural language without them. But of course that is the norm in modern societies. For instance, I have communicated in English via writing at length and on multiple occasions with various people who I have never met or talked to. Orthographies show that people without great linguistic sophistication have no trouble at all separating out languages from their original token forming mechanisms. That is, they substitute written token systems for spoken ones with hardly a thought.

It is notable how indifferent people are to the actual physical nature of tokens. When a document is read in a paper version, the tokens are some kind of ink marks on paper. When the same document is read on a computer or an e-reader or carved into stone, the physical nature of the tokens is different but one hardly notices. This underlines the distinctness of physical tokens from the types they are tokens of. More importantly, these alternative realization facts show the independence of the sentences themselves from any system associating them with physical tokens. The token producing mechanisms involved with handwriting, typewriting, computer printing, e-book displays, computer displays are all physically different. So the differences between realizations of a sentence S when printed on paper, displayed on a computer screen or on an e-reader or written in the sky by a skywriter indicate nothing about S itself.

A sharply contradictory view is seen in the following claim that it is necessary for grammatical statements to say something about concrete inscription tokens:

(4) Pullum and Scholz (1997: 43-44)

"Adequate principles defining syntactic structure must be satisfied by both abstract and concrete objects. For example, consider the generalization expressed within generalized phrase structure grammar as 'P < NP'. This linear precedence rule says, intuitively, that a member of the category P (a preposition) must not

follow any NP sister that it may have. The generalization it expresses is satisfied by abstract sentence structures, as platonist realists propose; for example, consider a class of binary phrase structure trees representing prepositional phrases that no one has ever uttered or been disposed to utter, nor ever will. Such a class is plausibly made up of abstract objects. Such a class satisfies 'P < NP' if and only if no tree in the class has a constituent with a left branch labelled NP and a right branch labelled P. (It matters not at all, incidentally, if some of the objects in the class are of infinite size by virtue of having infinite depth of embedding down some path.)

But the same statement, 'P < NP', is also true of all corpora of concrete English inscription tokens.This is why this statement- basically, the statement that English is prepositional - inclines linguists to think that inscription token sequences like (l la) are to be expected in tomorrow's newspaper, while sequences like (llb) are not.

(11) a. - a t - t h e - Wh i t e - H o u s e .

　　 b. - t h e - Wh i t e - H o u s e - a t .

The expectation is warranted; for example, not a single sequence of marks like(11b) appeared in the *Wall Street Journal* between 1987 and 1989 while sequences like (11a) showed up many times. When we assert that the statement 'P < NP' is about concrete inscription tokens as well as abstract grammatical structures, we mean that it is a condition of adequacy on theories of English grammar that they must say something about English inscription tokens, and they must say something about abstract constituent structure, grammatical relations, etc."

However, this argument is unsound because the premiss 'But the same statement, 'P < NP', is also true of all corpora of concrete English inscription tokens' is totally false. No concrete inscription can satisfy a claim about NPs and Ps. Examination of spectrograms will reveal no elements of any kind corresponding to formal representations of the grammatical categories NP or P or any others. Therefore, the question of their relative order in concrete phonetic inscriptions cannot arise. Just so for written tokens. Regardless of the medium, ink on paper, light patterns on computer screens, carvings in stone, etc., no physical analysis of tokens can reveal abstract entities like NP and P.

As for the data about the *Wall Street Journal* or any other token occurrences of sequences like those in the authors' (11), there is no mystery to be explained. Native speakers in general, outside of poetical contexts, usually write grammatical sequences, at least as far as gross grammatical constraints like preposition/nominal order. And in general, when constituent A precedes constituent B in a grammatical constituent C (an abstract object), in any token of C containing a representation of a construction [A B], the subtoken of A precedes the subtoken of B. This follows because the token forming mechanisms code a grammatical constraint A < B in speech as the instruction to create the token of A earlier than the token of B. And in Western language orthographies, a grammatical constraint A < B is coded as the requirement to write the token of A to the left of the token of B. Thus the facts cited by Pullum and Scholz follow from these more or less trivial considerations with no motivation to talk in addition (and incoherently) about a grammatical constraint being true of tokens or their parts.

Thus the argument in (4) shows nothing about natural language ontology and in no way supports the idea that languages involve mixed entities combining the properties of abstract objects and physical things.

It is in the context of the above conclusions that one should evaluate the claim of Postal (2004: Chapter 6, 192) to the effect that direct discourse sentences like (5) are such that the noises seemingly making up the object nominals of the verb *yelled* are actually parts of the sentences themselves.

(5)   a. Felicia yelled 'smekto'.

b. Felicia yelled 'smektof'

The claim was expressed as follows:

(6) Postal (2004: 192)

"The only way I see that this can be the case is if direct speech segments involve sets that contain the physical properties themselves and not, as in the case of more standard (regimented) linguistic elements, symbols that represent instructions (to a fixed physical apparatus) to produce physical things."

View (6) was motivated inter alia by the observation that there are endlessly many sentences like (5a, b), each with a distinct truth value, representing arguably distinct sentences. Claim, (6) was intended to provide a mechanism allowing for this class of distinct sentences. An alternative, namely, just positing some arbitrary symbol Δ and claiming it is the unique object of all direct discourse sentences like (5a, b) fails to differentiate the mass of cases with distinct meanings and truth conditions.

One might attempt to skirt this argument by observing that sentences like (5a, b) are indexical, justifying assuming that their truth conditions depend on pragmatic assumptions; see Schlenker (2017) for a recent discussion of indexicals. But taking all sentences like (5a, b) etc. to have the single object Δ crashes against the fact that (5a, b) represent distinct sentences, not a single sentence with contextually determined distinct truth conditions. That is justified by the fact that speakers recognize that e.g. (7a, c) are false.

(7) a. The number of distinct sentences represented by the distinct orthographic tokens in (b) is one:

    b. (i) Felicia yelled 'smekto'.

     (ii) Felicia yelled 'smektof'

    c. (bi) and (bii) are tokens of the same sentence.

Does claim (6) make literal sense in combination with previous claims in sections I-V? The answer is not entirely clear. Arguably there is no logical inconsistency between view (6) and my current claim that sentences are abstract objects. That follows because the 'physical things' in (6) were taken as elements of sets, so that even with physical

elements as atoms, sentences were still fully set-theoretical and hence abstract objects. The relevant sentences were just not taken be 'pure sets', those all of whose ancestral components are abstract, and could not, for example, exist in all possible worlds. While a consistent position, this is not really a comfortable view for a platonist account of natural language ontology. So the issue is whether on can today give a better account which, in particular, denies that physical properties have any status at all as parts of natural language sentences.

That there could well be a better account is plausible since one observes that (6) represents merely an argument from lack of imagination ('the only way I see'), which thus has almost no real force. As noted in (13) of Part IV, Pullum (2015: 15) took up this position of mine. But he gave no more argument than Postal (2004) did that sounds are parts of sentences.

Moreover, the view that token forming systems are not part of natural languages as such permits an arguably superior view as follows. Instead of taking physical sounds such as 'smekto' and 'smektof', to be parts of sentences, as in (6), one can now say that the terminal elements of the object nominals of *yelled* in the two cases are not noises or set theoretical structures whose atoms are noises but distinct *real numbers, N_1, N_2.* Taking N_1, N_2, etc. to be in general real numbers and not integers permits distinguishing the nonrecursively enumerable class of noises which can be referred to in examples like (6), hence distinguishing the nonrecursively enumerable class of sentences of the relevant form, each with a distinct truth condition. One then assumes that the denotational interpretation of the object DP whose terminal is some N should be characterized along the following lines:[25]

(8) The denotation in use X of sentence S of direct discourse DP_x whose terminal is real number is a noise Q, such that the the token of DP_x in use X of S reasonably imitates Q.

In full accord with a platonist view of natural language ontology, this revision maintains the type/token approach central to the present discussion, applying it to the particular class of cases like that in (5a, b). It makes a distinction between real numbers functioning as grammatical terminals and their physical tokens.

This proposal contextualizes the denotation of DP_x to the uses of sentences containing it in a bit like the way one must contextualize the denotation of first person elements to the uses of sentences containing them. That is, one cannot specify the denotation of e.g. *me* in any sentence containing it without referencing aspects of the contexts of the use of sentences containing *me*, specifically that aspect fixing the speaker. Just so one cannot say what the objects of the verb *yelled* in (5a, b) denote without referencing the contexts of use of the sentences in (5a, b) because only those contexts indicate the tokens which attempt to imitate some distinct performances.

The claim that the terminal string of the object nominal in sentences like (5a. b) is some real number eliminates any motivation from the kind of facts invoked in Postal (2004: chapter 6) and mentioned above for a view that such have physical sounds as set-theoretical atoms.

However, it might be an oversimplification to claim that DPs like the objects in (5a, b) are simply of the form of a real number. The question is whether such a view offers a proper ground for the calculation of the meaning of the containing verbal phrases from those of the verb and that of the object. In particular, either of (5a,b) entails (9):

(9) Felicia yelled something.

Given such entailments, perhaps the actual DP object in such a case involve a richer structure involving the elements of an existential quantifer phrase like that in (9) expanded by a kind of modifier phrase, as in:

 (10) Felicia yelled [something of which Δ is a reasonably accurate imitation]

In the direct discourse case, given any token T of (5a) with structure (10), the denotation of the object DP is determined by that subpart of T which is the token of Δ.

But the correctness of these structural speculations is irrelevant to the central ontological point which concerns what if any implications direct discourse sentences have for an overall view of natural language sentence ontology. And the overall conclusion is that if, as argued, the token forming mechanisms are not part of natural languages, then their properties cannot as such have a bearing on the ontology of natural language sentences or natural languages. If though direct discourse sentences implicate

the existence of real numbers as terminals, as I have suggested, then they do conribute to showing that natural languages are not recursively enumerable collections.

Incidentally, it is possible that the sharp division between natural languages and the token forming sysems associated with them I have advocated here overlaps with the views of Hjelmslev (1953); see Garvin (1954). But unfortunately, I am not familiar enough with his ideas to make a serious comment.

Part VII On the Concept *Mind-Dependent Abstract Object*

Pullum and Scholz (1997) take seriously the idea that linguistic objects can both be abstract and mind-dependent, that is, dependent on human minds for their existence. They suggest that such objects define an alternative both to fully naturalistic and to platonist views of natural language ontology, as in:

(1) Pullum and Scholz (1997: 30)

"It is not controversial to assume that an abstract object can be mind-dependent. Some medieval philosophers held that (as Quine 1948 puts it) 'there are universals but they are mind-made,' and that view has traditionally been called 'conceptualism'. They were not proposing that universals are physical objects located in brains, but did propose that they are constructed by cognitive activity."

Pullum (2015) criticizes Katz (1981) and Katz and Postal (1991), whose platonist advocacy provided no discussion of the possibility of mind-dependent abstract objects. I cannot speak for Katz, who died in 2002, but I regard our previous lack of such discussion as emminently defensible since I think that if the view that abstract objects can be mind-dependent is not controversial, it certainly ought to be. What follows is an effort to show why it is not only controversial, but worse, is an idea which has not been shown to be coherent, still less linguistically useful.

The view is advanced again in:

(2) Pullum (2015: 7)

"If Platonism about linguistic objects is the view that sentences (etc.) are abstract and mind-independent, then it has more than the two rivals (nominalism and conceptualism) that Katz discusses. At least the following possible positions concerning English Sentences conflict with Platonist realism:

(3) a. Katz's 'Nominalism': English sentences exist mind-independently but are concrete rather than abstract.

   b. Katz's 'Conceptualism': English sentences exist as mental representations: they are both concrete and mind-dependent.

c. Pullum & Scholz's 'Constructivism': English sentences exist as abstract yet mind-dependent objects."

Despite Pullum and Scholz's view (3c) that linguistic objects might be both abstract and mind-dependent, they provide no analyses of the concepts which would be needed to make such a view coherent, specifically *abstract* and *dependent*. Normally, as touched on earlier, abstract objects are taken to be instantiated by things like those cited in (3):

(3) Rosen (2017: 1)

"Some clear cases of abstracta are classes, propositions, concepts,

the letter 'A', and Dante's *Inferno*."

The entities cited do not exist in space or time and can neither cause nor be caused. It has been argued in earlier sections (and of course elsewhere) that natural language sentences are just such things. Pullum and Scholz say nothing contradicting this characterization. But at the same time, I find no clear statement that they accept it or any other specific characterization. I conclude that at this point, criticism of the notion mind-dependent abstract object probably should branch. On one understanding, authors using this term as if it denoted something have so far failed to characterize a graspable notion.

On the other branch, taking their notion mind-dependent abstact object to reference abstracta *as now usually understood*, a web of problems arise as to how objects abstract in this sense could depend on human minds (or any others for that matter). While I do not propose to discuss the profound question of the nature of minds, I strongly doubt that Pullum and Scholz were willing to contemplate a view of them such that they did not exist in time or space. But if minds have spatio-temporal coordinates, what notion of dependence could there be such that things not having spatio-temporal properties could depend on objects having them?

Remarkably again, I find no clear statement at all as to what the term *dependent* in *mind-dependent abstract object* is supposed to denote. Quote (1) assumes it is sensible to think of an abstract object as being 'constructed by cognitive activity', where this is not taken as mere metaphor. But the needed notion of *construction* is left obscure. I do not

see how to give it any coherent interpretation. What kind of things are abstract objects and cognitive activity taken to be such that the latter can create the former? If abstract objects are not found in time or space and cannot enter into causal relations, how can cognitive activity, which is surely temporal, spatial, and, at least in any naturalistic approach, part of nature, bring about their existence? Moreover, once the putative construction of a mind-dependent abstract object is complete, what kind of entity is supposed to have been created? If it is an abstract object tout court, why the special terminology? And if it is not, then no characterization of the supposed entities has been provided. The situation is parallel to that alluded to in Part I with respect to Chomsky's failure to provide any graspable notion of what natural language sentences are in his so-called biolinguistics. Pullum and Scholz have left these matters strikingly mysterious.

Consider a linguistic thing of interest, e.g. a particular sentence like (4):

(4) Most authorities consider time travel to be an indulgence of rich white people.

If this was created by cognitive activity, we are justified in asking whose. It should be troubling then to the advocate of mind-dependent abstracta that no creator of (4) is to be found. In fact, (4) has an impersonal existence, unlike, e.g. my view of the idea of mind-dependent abstract objects. I conclude that talk of the construction of abstracta by cognitive activity has no known meaning. Arguably then, branch two of criticism of the Pullum/Scholz proposal about mind dependent abstract objects is, because of its lack of clarity, not so different from branch one.

Still another perspective on the notion of mind-dependent abstract things is this. A traditional problem for platonist views, not infrequently taken as the basis for a decisive *rejection* of such views, is the epistemological problem. That is, how can humans existing in time and space gain knowledge of things which do not and which do not enter into causal relations.[26] This criticism of platonist views assumes a particular view of epistemology, one requiring that noninnate knowledge of X can arise only via some kind of causal connection of the knower with X. But this rigid view of the origin of knowledge should not be taken as having been demonstrated. For instance, the discussion

in Wetzel (2009: chapter 2) provides ground for considerable skepticism as to this epistemological view, citing a range of earlier work supporting such skepticism.

Pullum seemingly accepts himself that the epistemological problem is a serious threat to platonist ideas:

(5) Pullum (2015: 8)

"But it is well motivated by worries of an epistemological nature: for objects that are not spatiotemporally located and that have no causal relations to us or our world, it is completely obscure how we could come to know about them at all."

Ironically though, setting aside the lack of clarity in the component ideas, if there is an epistemological problem, the idea of mind-dependent abstract objects appears to incorporate its essence. If it is a mystery how human minds could gain knowledge of things not existing in time and space, and with which causal relations are impossible, it is surely an equal or greater mystery how abstract things can *depend for their existence* on human minds. In both cases, the problem is how any relation can exist between causally functioning creatures existing in time and space and objects not part of the physical world.

One can well fear at this point that talk of mind-dependent abstract objects appeals to some notion 'abstract' which does not denote what is usually meant by that word. That would be unobjectionable if the proponents of the view had given a clear analysis of what it *is* supposed to denote. But Pullum and Scholz do not address these issues. Given that, I am inclined to conclude that their claims that languages are mind-dependent linguistic objects, to the extent they are not simply incoherent, have not been elaborated sufficiently so far to permit questions of its truth to arise.

Pullum and Scholz provide no serious analysis of any actual linguistic object about which it would make sense to say both that it is an abstract object and yet dependent for its existence on human minds. One does find:

(6) Pullum (2015: 8-9)

a. "It is surely not irrational to suggest that if there were no thinking beings in the universe there would be no sentences of English or any language.

b. "It seems to me comparable to the notion that if there were no thinking beings in the universe there would be no money." "We regard a diamond deep underground that no human has yet discovered as existing: no human being needs to know about it or think about it in order for it to exist. Nothing similar holds for a dream that no one has ever had, or for a poem that no one has yet written, or a sentence that no one has ever thought to frame or utter. I think it is reasonable to regard such things as dreams, poems, and sentences as abstract yet mind-dependent."

Talk of irrationality here is a bit of a straw man as it is possible to be perfectly rational and mistaken.

Consider the supposed examples in (6) one by one. Pullum wrote *a dream that no one has ever had*. By his own philosophy, this phrase cannot denote anything. But in reality it can. The reason is that like other terms such as *book*, see below, there is an ambiguity. One one reading, dream DPs denote mental events. Consider:

(7) I had a dream that Arthur was a space alien and Tom had the same dream.

On the mental event reading, (7) is incoherent (under the assumption that no two people can share the same mental event). On event readings, dreams take place in space and time:

(8) Last night, when I was in my room in Vancouver, I dreamt that I was chased by two vampires.

But on another reading, a dream DP refers to a kind of scenario. Then (7) makes perfect sense, claiming not that Tom and I experienced the same mental event but rather that our distinct mental events instantiated the same scenario, in the way that two physical recordings or playings can instantiate the same film. In this sense, there are endless dreams which no one has had and endlessly many which no one can have (because of length/complexity considerations, etc.). The scenarios here would be abstract objects, while the mental events would be mind-dependent. But neither would represent mind-dependent abstract objects.

Pullum's second reference to things putatively both mental and abstract picks out poems, say literary objects in general. I discuss these in detail below. Third, Pullum suggests that natural language sentences only exist as a result of humans having framed or uttered them. I confess to a lack of clarity as to what framing a sentence involves. But no matter; clearly, it references something mental.

Consider:

(9) a. No six-headed space alien lizard spoke very harshly to Ted.

b. No six-headed space alien lizard spoke very very harshly to Ted.

c. No six-headed space alien lizard spoke very very* harshly to Ted.

Here, according to Pullum we can take (9a, b) to represent existent English sentences since I just formulated them. Now take * to be the Kleene star schematism representing 1 to n sequences of the adverb *very*. Utilization of that schematism would in Pullum's terms frame one or more other sentences which would exist only because of that framing. But, assuming no one has framed them before, they don't actually exist as yet. But hey, let's make another one exist:

(10) No six-headed space alien lizard spoke very very very harshly to Ted.

There I did it, or at least did it if one takes Pullum's view seriously.

Beyond being bizarre, Pullum's view is, notably, thoroughly incompatible with his adoption of a model-theoretic view of natural language grammars, described in various of his papers over the last decade or more, in particular, as follows:

(11) Pullum (2013:49-498)

a. All rules are constraints — truth-eligible statements about the internal structures of expressions.

b. Grammars are unordered sets of such constraints (theories, in the logician's sense).

c. Well-formedness of a structure is determined by satisfaction of the constraints.

This conception of grammars, which I endorse and which Johnson and Postal (1980) is one origin of, is strongly motivated. But (11c) in particular is entirely independent of whether a human has framed one or more such objects.

For instance, while it is vanishingly implausible that anyone has framed any of (9a, b, c) before, it is no more plausible that anyone has previously framed the endless number of other sentences entirely parallel to these where *six-headed* is replaced by some some other numeral expression + *head*. According to Pullum's advocated model-theoretic conception of grammar, all of these will surely be sentences of English. That follows since they could fail to be sentences only if there are constraints built into the grammar distinguishing e.g. *six-headed* from all of *nine-headed*, *forty-thre*e *headed*, etc., which is unmotivated and absurd. But according to the mind-dependent view, none of them can be. This has to be a reductio of either the model-theoretic view or the view that sentences get created (in some unknown way) by human mental activities. Since the former is eminently sensible and insightful, the conclusion is clear.

Independently of these considerations, I cannot see how the investigation of natural language sentences could procede under a view that that their existence is dependent on human framing or uttering. How can a student of English know what sentences have been previously framed or uttered (and thus are granted existence in this philosophy), since this framing can take place in untold numbers of minds? In practice, this would reduce linguistics to the study of existent texts and/or to the view that research involves creating the objects to be characterized.

There is a distinct issue. If natural language sentence existence depends on human framing or uttering, then the class of sentences in any language has to be finite because humans are finite creatures along every dimension, including the temporal one. In other words, Pullum's interpretation of his mind-dependence claims lead to the 'discovery' that there is a class of longest sentences. Were that so, it would be an important discovery about language and significant that it had been shown to follow from deep ontological assumptions. But the Pullum/Scholz view is that size considerations are, as noted earlier, outside the concern of grammar, a view stated strongly in:

(12) Pullum (2013: 502)

> "A typical MTS grammar would therefore not say anything about how
>
> big expressions can be. (One could stipulate, if one wished, that there are not

more than 37 nodes; a tedious and lengthy (but straightforward) first-order formula could do it. But absolutely no one thinks that would be a good idea, for 37 or any other number.) There is no need to ask whether we should regard billion-word sentences as grammatical but impossible for a human being to process: the grammar makes statements about what structural conditions must be respected, but it can be entirely agnostic on size."

However, if a 'typical MTS grammar' is embedded in an ontology like (1), the size neutrality of any such grammar is irrelevant as the underlying framework has told us that languages are finite and that sentence size is controlled by the finite limits on human nature.

Given the overwhelming weaknesses in an ontological view appealing to mind-dependent abstract objects, it would perhaps be best to say no more about the notion. However, probably Pullum/Scholz felt confortable in their view because it seems to a significant extent not to be original and one can cite past scholars who seem to have viewed the idea as reasonable.

For instance, Pullum (2015: 3-4) cites the following passage from the respected logician George Boolos.

(13) Boolos (1998: 128-129)

"We do not engage in physical interactions with them, in which energy is transmitted, or whatever. But we twentieth-century city dwellers deal with abstract objects *all the time*. We note with horror our *bank balances*. We listen to *radio programs*: *All Things Considered* is an abstract object. We read or write *reviews of books* and are depressed by *newspaper articles*. Some of us write *pieces of software*. Some of us compose *poems* or *palindromes*. We correct *mistakes*. And we draw *triangles* in the sand or on the board. Moreover, bank balances, reviews, palindromes, and triangles are "given" to us "in experience," whatever it may mean to say that. . . . [No] sense of "sensible" or "experience" has been shown to exist under which it is not correct to say that we can have sensible experience of such objects, such things as the zither melody in *Tales from*

the Vienna Woods, the front page of the sports section of the morning's *Globe*, a broad grin, or a proof in set theory. . ."

The citation is apparently meant to provide some support for the mind-dependent abstract object view. But Pullum builds no argument on the quote, which itself contains none.

What it does contain is just a failure to distinguish between certain abstract objects and things which serve as tokens of them. For instance, Boolos says people listen to the radio program *All Things Considered*, which he takes to be an abstract object. He is surely right on one sense of the expression. On that sense, the program cannot be found in space or time. But in that sense, one cannot listen to it. One listens to particular physical events, which arguably bear a token relation to some abstract object appearing under the aegis of *All Things Considered*, a still more general abstract object. The situation is analogous to that of dreams, discussed in connection with (6) and (7) above.

Similarly, despite Boolos' remark to the contrary, one cannot take various patterns of sand to be triangles. The latter are indestructible geometrical abstract objects, while the former physical things cease to exist at the whim of winds and tides. The other items cited likewise fail to distinguish abstractions from various sorts of concreta.

One can also cite current philosophical work which also appears to take a position compatible with the Pullum/Scholz view.[27] Consider the following remarks from the article Abstract objects at the *Stanford Encyclopedia of Philosophy*.

(14) Rosen (2017)

"A similar problem arises for so-called *abstract artifacts*, like Jane Austen's novels and the characters that inhabit them. Some philosophers regard such items as eternally existing abstract entities that worldly authors merely 'describe' or 'encode' but do not create. But of course the commonsensical view is that Austen created *Pride and Prejudice* and Elizabeth Bennett, and there is no good reason to deny this (Thomasson 1999; cf. Sainsbury 2009; see also the entry on fiction). If we take this commonsensical approach, there will be a clear sense in which these items depend for their existence on Austen's mental activity, and perhaps on the mental activity of subsequent readers. These items may not count as mind-

dependent in either of the senses canvassed above, since *Pride and Prejudice* can presumably exist at a time even if no one happens to be thinking at that time. (If the world took a brief collective nap, *Pride and Prejudice* would not pop out of existence.) But they are obviously mind-dependent in some not-merely-causal sense. And yet they are still presumably abstract objects."

This discussion of the specific example *Pride and Prejudice* permits seeing that something renders plausible the idea that e.g. written works are created by humans. Recall the phrase in Pullum and Scholz's (1) 'created by cognitive activity'. And yet it is easy to argue that they are abstract objects, as the author of (14) agrees, and distinguishable from their physical tokens. While the former do not exist in time or space, and are indestructible, the latter can be ripped up, burned, etc. And supposed things which are abstract but somehow created by humans are arguably exactly the sort of putative mind-dependent abstract objects Pullum, Scholz and Rosen are talking about.

But the specificity in the Encyclopedia work permits showing that the view is superficial and confused. My contrary view is that *Pride and Prejudice*, like *every other document ever*, is an abstract object. That is, like its component sentential expressions, it exists neither in time nor in space and was neither caused nor could ever cause anything. As such, its existence cannot possibly be dependent on human minds or their cognitive activity.

It is thus important to delve into the writer's claim that there is no reason to deny the commonsensical view that Austen created *Pride and Prejudice* and his claim that objects like *Pride and Prejudice* are *obviously* mind-dependent. Since I am not familiar with that august work, let us take a shorter and simpler one, a short story that in exactly the commonsensical sense of Rosen (2017) I now create.

(15) Short Story

The space alien lizard ship landed in Central Park on a hot summer evening. The two lizard occupants of the ship exited it around 1AM and were immediately attacked, killed and stripped of all of their equipment by typical nightime denizens of the park. By 6AM, the space ship into which the lizard corpses had

been thrown had been thoroughly sacked and set on fire. As the materials of the ship were highly flammable, by 7AM, there was nothing whatever left of the ship or the lizards, not even chemical residues. This is why news reports give no indication that a space alien lizard ship ever landed in Central Park.

It seems undeniable that *if Pride and Prejudice* is to be taken as a mind-dependent abstract object, one created by Jane Austen, then (15) must be taken to be a mind-dependent abstract object created by me. But while there is indeed an abstract object *connected to* (15), it is *an illusion* that that abstract object is mind-dependent or was created by me. I did though create *something*, so the whole constellation of realities involves both abstract object and creativity, just not creation of the abstract object.

First, speaking commonsensically, (15) consists of the orthographic representation of five English sentences. A first idea might then be that the orthographic sentence tokens in (15) represent a set of sentences. But that cannot capture the nature of the story because the representations occur in a fixed order and clearly different orders of the representations correspond to different stories.[28] So stories are not sets of sentences.

A better idea then is that the story (15) represents consists of a set of pairs {n, x}, where x ranges over the five English sentences represented by the given orthographic elements and n ranges over a (for simplicity, continuous) set of integers from 1 to 5.[29]

If so, then since sets neither occur in time or space nor enter into causal relations, the set-theoretical nature of the story justifies taking it to be an abstract object. Is there though any sense in which its existence is mind-dependent and in which I created it? There is not. But the commonsensical idea that I created the relevant abstract object, and thus that in that sense it is mind-dependent, has, one must recognize, a kind of cloudy plausibility.

That is true, I believe, because the possibility of reading the discourse in (15) was dependent on my doing something. Take the standpoint of my cat Milo staring at me at the computer. What Milo would have seen, or a functioning video camera would have recorded, was me pushing the keys of a computer keyboard resulting in the appearance on the screen of certain marks. Neither Milo nor I have any idea of how pushing the keys

yields the marks on the screen. But he doesn't care and for my purposes I don't have to either. Milo is also unaware of what mental activities went on in me as I pushed the keys and is again indifferent. I have some vague idea about those activities. But whatever went on in my mind, the graspable real-world consequence of my creativity is just the orthographic representations in (15).

So the obvious question is what is the relation between a document-type abstract object and the production of orthographical marks on a computer screen, or a piece of paper. That is, what creates the illusion that one can create an abstract object? The clue to this puzzle is that any reader's idea that the successive marks in (15) represent a story about space alien lizards is entirely dependent on *the reader's knowledge of English* and English orthography.

Suppose instead I had presented:

(16) Tsi nikaianerenhseroten: thé:non ne ohwista enkontiié:na ne Kontinónhstats iáh thaón:ton ne ieià:tare thé:non áhontste ia'teieiá:ti aiakoianeráhsten.

Most readers will not find any story or statement in this collection of marks and will not even be able to tell whether it represents actual natural language orthographic representations or is simply at least partially a random collection of characters. In fact, it is *part* of a document in the North American Iroquoian language Mohawk.[30]

The answer then as to how creating a series of marks in some medium can relate to, essentially pick out, an abstract object is as follows. This result is accomplished by virtue of the fact that some such sequences of marks correspond to the visual tokens of natural language sentences according to some existing orthographic conventions. If, as in the case of (15), the marks represent orthographic tokens of *English* sentences, then someone *with knowledge of English and of English orthography* can interpret these marks as tokens of the relevant sentences and grasp that there is an object corresponding to the relevant set of integer/sentence pairs. In one sense, that object will not be the set of integer/sentence pairs but rather a set of pairs of integers and the *meanings* of the perceived sentences. That is why the notion of translation makes sense in principle. Because an *ideal* translation (or paraphrase), rarely if ever achieved, provides a distinct

set of integer/sentence pairs where, however, the corresponding set of integer/meaning pairs is identical to the original.

Thus to understand the relation between making marks in some physical medium and the existence of an abstract object which is a story, a novel, etc., no appeal at all is necessary to any mental creation of abstract objects or any dependence of the latter on minds. Several things other than the creation of the marks do of course depend on minds.

First, a mind-possessing creature is needed create the relevant orthographic marks.[31] Second, it is linguistic knowledge which yields the conclusion that a collection of physical marks like (15) represent tokens of the expressions of a language. Only that linguistic knowledge permits recognition that the marks represent a novel, story, poem, financial report, etc. But such linguistic knowledge is uncontroversially mind-dependent.

The just given account thus grants that there is a human creative process involved in the writing of books, articles, etc, while denying that such processes create abstract objects. Rather, what they accomplish is just to point out to readers equipped with the relevant linguistic knowledge the existence of certain necessarily uncreated abstract objects. In effect, authorship of a particular text T, through the intermediary of knowledge of a natural language and its orthography, picks out from the endless array of eternal abstract objects formable from sequences of natural language sentences (plus frequently things like those mentioned in note 29), a particular abstract object composed of ordered expressions represented by the ordered orthographical elements of T.

Put differently, authorship of some text T aids in the creation of knowledge of a particular complex abstract object. In principle, no one having knowledge of the language in which T is written needed an author to gain knowledge of T. But given the infinitude of texts formed by combinations of natural language sentences, in actual fact, no one has any chance to focus on T a priori because of the endless universe of existing alternatives. So in real life, writing texts serves as a catalyst for otherwise unobtainable knowledge of abstract objects.

The account I have given of document authorship hugs quite closely the observable physical facts. This view talks about the creation of physical marks and their

interpretation by physical creatures on the basis of their linguistic knowledge. It creates no threatened incoherence about the relations between physical and abstract things. It is perfectly compatible with common sense understandings. But it does not, of course, eliminate or in any way illuminate the mystery of how linguistic knowledge of the relevant abstract objects is possible or originates. Those are issues about which I have nothing to say. But since no one seems to deny that there is linguistic knowledge, it is unnecessary for present purposes to confront that problem.

Pullum and Scholz (1997) gave another type of argument intended to show that natural language sentences involve mental features and are thus mind-dependent objects, yielding a purported ground for the view that a platonist conception of language is untenable. It was based on supposed properties of the English Heavy NP Shift phenomenon:

(17) Pullum and Scholz (1997: 41-42)

"Every adequate theory of English syntax must refer to what are generally called 'heavy NPs': noun phrases eligible to participate in the Heavy NP Shift construction in English. This is illustrated in (9). The normal order of direct object and indirect object is shown in (9a). Reversing the two in this case yields the ungrammatical (9b). We would expect from this that (9c) would be grammatical and (9d) would not; but this is not the case. If anything, (9d), which illustrates the Heavy NP Shift construction, is more acceptable than (9c).

(9) a. They gave those to each student.

b.*They gave to each student those.

c. They gave a large scroll on which were written some words in Latin to each student.

d. They gave to each student a large scroll on which were written some words

in Latin.

The property that makes a noun phrase heavy is not a structural property like having a subordinate clause contained in the NP, as shown by (10a), where there is no such subordinate clause but Heavy NP Shift is permitted; nor is it concrete, like phonetic length (milliseconds of time taken to utter it), as (10b) shows. hey gave to each student a large parchment

(10) a. They gave to each student a large parchment scroll.

b. They gave to each student ... one of these.

Yet as we saw in (9b), not every NP can count as heavy. So what makes an NP heavy? The answer is that it is heavy if it is judged heavy. Certain qualities - length, syntactic-complexity; prominence, pragmatic surprisingness, 'newness' of the information conveyed-may influence a speaker to treat an NP as heavy enough to be shifted to the end of the verb phrase (assisting the hearer in processing it) but no properties independent of the speaker's judgment determine heaviness. There is no difference between seeming heavy and being heavy. The representation and the object are the same thing. Thus a heavy NP is a mind-dependent object in the second sense we discussed under (1b.ii)."

However, this argument about the Heavy NP Shift construction is devoid of content, signaled by the fact that it states no coherent constraint defining the phenomenon. Basically, the logic of (17) is just the following, with the nonsequitur (18b):
(18)   a. Premiss: The authors are unable to state any coherent condition distinguishing permitted from barred instances of Heavy NP Shift.

b. Therefore, there is no coherent nonmental condition on Heavy NP Shift.
That is, the authors try to turn the lack of insight represented by (18a) into a fact about language. But the same awful reasoning would permit any authors to turn any failure to

understand the conditions on any aspect of language into a conclusion that there are no statable conditions. Moreover, this dreadful embrace of failure is based on the analysis of only the tiny data base in their (9) and (10), containing just four actual instances of Heavy NP Shift, their (9a, d) and (10a,b).

And once one corrects what I find to be an equivocation in their proffered data, there is a statable and even quite simple constraint on the construction, namely, informally, the following:

(19) If a nominal phrase shifted under Heavy NP Shift consists of a single word W,

W is strongly accented.

If we orthographically mark strong accent with caps and its absence in shifted phrases with lower case, then (19) combines with independent constraints on the possibility of strong accent to block cases like (20a, b, c) while allowing (20d):

(20) a. *They sent to the children it.

b.*They showed to the victim her.

c. *They sent to the children IT.

d. They showed to the victim HER.

The difference between (20c, d) is just that *it* seems to *generally* preclude strong accent, everywhere independently of Heavy NP Shift, while *her* does not.

Constraint (19) allows unproblematically the authors' (9d) and (10a, b), each with a shifted multiword nominal. So the only potential problem for (19) in Pullum and Scholz's data set is their (9b), which they star. But this is just a *potential* problem since Pullum and Scholz do not discuss the accent on *those* and I find the example grammatical when *those* is strongly stressed. So the claim that 'no properties independent of the speaker's judgment determine heaviness' is groundless. Moreover, if the need for strong accent is marked on a subconstituent by some grammatical marker, there is no need whatever to conclude that (19) needs to refer to the phonetic properties involved in strong accent.

Condition (19) is probably a consequence of a more general condition. That is, even multiword nominals which successfully manifest Heavy NP Shift must, I believe, manifest strong accent. So compare:

(21) a. Sylvia showed the book to the children.

    b. *Sylvia showed to the children ... the book.

    c. Sylvia showed to the children ... the BOOK.

The shifted phrase in (21b, c) cannot be accented like the unshifted phrase in (21a).

Overall then, the discussion of Heavy NP Shift in Pullum and Scholz (1997) offers nothing in the way of support for the existence or linguistic relevance of mind-dependent abstract objects.

One conclusion then is that talk of mind-dependent abstract objects appears to be entirely otiose. Even if (counterfactually, as I have argued) there were a clear coherent idea of what such things could be, appeal to them seeks to solve problems whose solutions have no need for them. Minimal Okham's razor considerations thus weigh heavily toward the conclusion that the idea they have something to contribute to the understanding of linguistic (or other) reality is a mistake. There is no ground for believing that they could provide an alternative to other conceptions of natural language ontology, specifically the platonist view.

Part VIII The Bottom Line

I have no doubt that now and in the past linguists have almost universally believed that natural languages are a natural world phenomenon whose existence is inherently linked to human nature. Chomsky's decades long talk of linguistic faculties, linguistic organs, langauge innateness, etc., have solidified this assumption and given it a special formulation. Key to the view is a failure to draw any distinction between natural language(s) and knowledge of natural language(s), advanced inter alia with unsupported, question-begging pieces of dogma like:

(1) Chomsky (1972: 169 n3)

> "Since the language has no objective existence apart from its mental representation, we need not distinguish between 'system of beliefs' and 'knowledge' in this case."

See Postal (2012) for fuller discussion of claims like (1). But once one distinguishes between a language and knowledge of it, as in (21) of Part III, as one must given their radically different properties, the uncontroversial human character of the latter provides no ground as such for associating these properties with the former.

The Pullum/Scholz appeal to ideas represented by the words *mind-dependent abstract object* deviates from a simple physicalist view of natural language by seeming to grant some abstractness to natural language sentences. But nothing has been said to show how a coherent formulation can be given of such a putatively mixed view.

Since Pullum and Scholz do grant *in part* the abstract nature of natural language sentences, one is led to ask why they retain a belief that sentences are also in some unknown way mental/human. I am convinced that this has two bases. First, there is their failure to distinguish language from knowledge of language, found as much in Pullum's recent writings as in Chomsky's and sharply revealed by the discussion in Part V. The mental, perhaps physical, nature of knowledge of natural language then leads, if the distinction is not made, to incorrectly attributing defining properties of linguistic knowledge to languages themselves.

The second ground for their failure to adopt a purely abstract object view of natural language sentences is the traditional assumption that the preorthographic token forming system of a natural language (phonetics in the standard case) is an inherent part of it. But that view, despite its past and present near universal acceptance, seems to offer no insight into anything, no more than would the idea that printer technology is part of Microsoft WORD. Having rejected that as well, I see no linguistic grounds at all for regarding natural languages as anything but abstract objects completely devoid of any physical or mental elements or characteristics.

To conclude, it is important that a platonist view of natural languages, one which by definition takes them to be abstract objects and only abstract objects, in no way leads to conclusions like those in the following:

(2) Pullum (2015: 10)

> "Pullum & Scholz (1997) aimed to open up some space for two further views: heterogenism, which says that more than one ontological variety of stuff is relevant when interpreting theories of linguistic structure, and pluralism, which says that separate and distinct homogeneous domains of discourse can participate in making the statements of a linguistic theory true.

> Their thesis is that even if there is a platonic realm where sentences exist independently of time or space, that doesn't mean that linguistics has to talk about that domain and no other. Katz, Postal, Landendoen, and others are in danger of propounding a linguistics that, by the very definition of its foundations, has absolutely nothing to say about utterances that are actually attested in written or spoken form, or about the mental operations that only a language-using animal like us can perform."

The last sentence of this passage is just wrong. A platonist view of the nature of natural language sentences and the collections of them forming languages in no way requires defining a field which excludes the study of token mechanisms and the resulting tokens

or language learning or language processing or the historical facts about language, etc. etc. This is nothing more than a question of the definition of *linguistics* and nothing in a platonic view requires any new one. The danger Pullum alludes to is thus not real since the existing usage of the term in no way excludes the things he cites.

Notes

**I am indebted to David E. Johnson for many comments on an earlier version of the manuscript which have significantly improved this version.

1 This view of abstract objects seems relatively standard in philosophy; see e.g. Cowling (2017: 2). See Katz (1981, 1984, 1996), Katz and Postal (1991) for application of this view to natural language in particular. Part VII considers claims that there is at least a third type of thing natural language sentences might be, distinct from being purely abstract things or straightforward exemplars of physical things.

2 The latter include Langendoen and Postal (1984, especially chapter 6), Postal (2003, 2004, 2009, 2012). Except for Postal (2004), the cited articles are available at Lingbuzz.

3 See for instance Chomsky (2005: 1).

4 For a description of Hilbert's Hotel see e.g. Rucker (1995: 73-75).

5 See http://itre.cis.upenn.edu/~myl/languagelog/archives/003086.html

6 This section is a trivially edited version of the following LANGUAGELOG post:

http://itre.cis.upenn.edu/~myl/languagelog/archives/003109.html

I am indebted to David Beaver for originally posting it and for recently helping me recover the reference.

7 It would, to say the least, be an odd organ that had no specifiable physical properties. But worse, NC's assumptions do not permit the hypothetical organ to have physical properties, since he claims *both* that a human language is a state of the innate organ and is discretely infinite. Nothing infinite can be a physical organ or a state of such, on the naturalist assumption, which NC of course makes, that human bodies are physical. Hence the claim that the posited organ has something to do with biology is not serious. Real organs, e.g. livers, are all too finite. Moreover, real organs can't produce an infinite number of, or amount of, anything, e.g. bile. The bottom line is that NC's position that natural language is both to be taken as an organ state and as discretely infinite is simply incoherent.

8 The syntactic approach which comes to mind involves an idea little mentioned these days which appeals to something called *conjunction reduction*. On that rarely made precise view, Seuren's example is a reduction of:

(i) John wants to get married and every woman in the village wants to get married.

Here the quantifier DP *every woman* scopes in the clause making up the second conjunct, hence not over a conjunct or a predication involving *John*. One can recognize the great appeal of the idea of conjunction reduction while granting that no serious account of it seems to exist internal to any currently advocated conception of natural language grammars.

9 On the relevant conjecture, see:

https://en.wikipedia.org/wiki/Goldbach%27s_conjecture

10 On the Continuum Hypothesis, see

https://en.wikipedia.org/wiki/Continuum_hypothesis

11 For instance, one question answer pair was:

(i) QUESTION: Is it true that language is a superstructure on the base?

ANSWER: No, it is not true.

Another was:

(ii) QUESTION: Is it true that language always was and is class language, that there is no such thing as language which is the single and common language of a society, a non-class language common to the whole people.

ANSWER: No, it is not true.

I agree with comrade Vissarionovich on these points.

12 See Chomsky (2012: 91), and the detailed discussion in Postal (2012). This one time event should not obscure, however, that Chomsky has otherwise ignored his admission and continued to assert his incoherent position, ignoring works pointing it out. Not exactly a model of intellectual courage.

13 This view is discussed in Part VII. The cited authors also contemplate still other ontological positions about natural languages, which I will not try to discuss here.

14 This hypothetical does not require acceptance that there is such a possible world.

15 This claim presupposes the view of Part V that the articulatory movements and resulting sounds used in most languages to create tokens of natural language sentences are not parts of the languages themselves.

16 In a framework where natural language grammars are regarded as model-theoretic systems, as in inter alia Johnson and Postal (1980), Postal (2012), Pullum (2007, 2013) and Pullum and Scholz (2001, 2005), one is led to describe the case of 'new' words like *Brexit* in a nonconventional way. The idea is that the phonological structure (an abstract object) defining e.g. *Brexit* existed in earlier forms of English as well as those postdating Great Britain's membership in the European Union and discussion of its exit from that system. The difference between the languages existing before and after Britain's leaving the European Union became an issue is as follows. *Brexit* would involve an earlier grammar in which this form was useless because, in the absence of specific constraints, it was associated with every possible grammatical categorization and *every possible meaning*, including e.g. 'hate vegetarians' and 'chocolate pudding'. The newer language restricts the meaning to a certain proposed event. This view takes every phonologically licit phonological combination in a language, e.g. *vestom* in English, to be an element of the language, although most of them are in practice unusable because of their length and their syntactic and infinite semantic ambiguity. See the discussion in Johnson and Postal (1980: 675-677), Pullum and Scholz (2001: 33-34) and Pullum (2015: 15-17).

17 This is an oversimplification as there is no reason to assume even for very small human communites that at a fixed time all the adult members had knowledge of exactly the same natural language. But the oversimplification is harmless in the present context.

18 The performance limitations on the production or processing of sentence *tokens* have nothing to do with specifying N. For support for this view, see Postal (2003).

19 Within the framework of Langendoen and Postal (1984) this remark needs to be hedged. Given the argument there that the totality of the sentences in English corresponds to the endlessly many ever greater power sets which are entailed by Cantor's Theorem

(see below in the text), technically the size of the whole domain of sentences is too great to be assigned any cardinality. One might even put it that there is no totality.

20 Were Pullum's skeptical view of the infinitude of natural languages well-grounded, it would have to have an analog for integers. One denying that there are infinitely many natural numbers would be commited to some N giving the largest, with no known basis for any choice.

The parallel between sentences and integers is hardly surprising since, excluding sentences of the type discussed in Postal (2004: chapter 6), e.g. direct discourse sentences like that mentioned in (13) of Part IV, all natural language sentences based on a fixed finite vocabulary could be coded as integers, via Gödel numbering; see https://en.wikipedia.org/wiki/Gödel_numbering. But reasoning about arbitrary choices of N again of course suggests that doubts about the infinitude of the natural numbers are unfounded.

21 One finds the following claim:

  (i) Pullum (2013: 496)

  "No evidence supports the claim that English couldn't have an infinitely long sentence,
   or could tell against it; it is simply stipulated."

However one evaluates the argument of Langendoen and Postal (1984), it did not *stipulate* the existence of nonfinite sentences but derived their existence from the principle of coordinate closure. See the No Upper Bound Theorem (p. 65) in that work.

22 Aristotle had already pointed out objections to finitism:

  (i) — Aristotle, Physics, Book 3, Chapter 6, cited at:
https://www.wikizero.com/en/Finitism

> "But on the other hand to suppose that the infinite does not exist in any way leads obviously to many impossible consequences: there will be a beginning and end of time, a magnitude will not be divisible into magnitudes, number will not be infinite. If, then, in view of the above considerations, neither alternative seems possible, an arbiter must be called in."

The Pullum/Scholz question-begging criticism thus in effect demanded that Langendoen and Postal enter and resolve in favor of the nonfinitist position a debate which has been going on at some level for at least two thousand three hundred years.

23 Although Stoll (1963: 1-2) does point out that Cantor's work on the infinite was subject to rejection by some mathematicians. He attributes this to prejudices.

24 Incidentally, Pullum's calling nonattested languages 'invented' here, as his earlier reference to them in (2) as 'imaginary', is unjustified persuasive terminology. It begs the question platonism raises, since for the platonist, abstract objects exist, can be discovered or not, but can never be invented and certainly from that perspective unlearned natural languages are no more imaginary than e.g. numerical calculations which have never been made. In rejecting the use of terms like *invented* and *imaginary* here, I am not presupposing the correctness of a platonist view. The objection is to a purely terminological, sustantively unsupported denigration of that view.

25 I am indebted to David E. Johnson for suggesting a formulation like (8) of Part VI.

26 For instance, the putative problem is formulated as follows by a scholar known for rejecting the existence of abstract objects:

> (i) Field (1989: 68)
>
> "The truth values of our mathematical assertions depend on facts involving platonic entities that reside in a realm outside of space-time. There are no causal connections between the entities in the platonic realm and ourselves; how then can we have any knowledge of what is going on in that realm? And perhaps more fundamentally, what could make a particular word like 'two', or a particular belief state of our brains, *stand for* or *be about* a particular one of the absolute infinity of objects in that realm?"

27 See also Cowling (2017: 5), who remarks:

> (i) "For expansive platonists, soup recipes, musical works, and a motley assortment of other entities are properly counted as abstract, despite their apparent dependence on human activities and their apparent location in space or time."

28 Also, since sets are defined by their membership, the view that written works consist of sets of expressions tout court entails that it is impossible for a document to contain the same sentence in several places, which is evidently wrong.

29 In general, this has to be a simplification, since written works can contain nonlinguistic elements such as pictures, diagrams, tables, mathematical and logical symbols, arbitrary marks, etc. Developing a more adequate account is not necessarily trivial but I do not believe these matters bear on the ontological points at issue here.

30 The relevant document is found at:

http://www.kanehsatakevoices.com

The English translation of the whole is:

> (i) "The Association was incorporated under Part II of the Canada Corporations Act by letters patent August 12, 1999. The association is a not-for-profit corporation, which means that none of its members may benefit personally from the association. This association seeks funding and donations from different foundations to achieve its goals for the preservation, revitalization and teaching of the Mohawk language, culture and history."

31 Well, not exactly. There is the *infinite monkey theorem:*

> (i) https://en.wikipedia.org/wiki/Infinite_monkey_theorem
>
> "that a monkey hitting keys at *random* on a typewriter keyboard for an <u>infinite</u> amount of time will almost surely type a given text, such as the complete works of William Shakespeare. In fact the monkey would almost surely type every possible finite text an infinite number of times. However, the probability that monkeys filling the observable universe would type a complete work such as Shakespeare's *Hamlet* is so tiny that the chance of it occurring during a period of time hundreds of thousands of orders of magnitude longer than the age of the universe is *extremely* low (but technically not zero)."

Since the relevant monkey typist will understand nothing of his output, this thought experiment stresses in another way that the creation involved in writing terminates in nothing more than the resulting orthographic objects.

References

Atkinson, Martin, David Kilby, and Iggy Roca. 1982. *Foundations of general linguistics*. London: George Allen and Unwin.

Beaver, David. 2006. Can you speak in Rhinoceros? http://itre.cis.upenn.edu/~myl/languagelog/archives/003086.html

Behme, Christina and Martin Neef (eds.). 2018. *Essays on linguistic realism*. Amsterdam: John Benjamins.

Boolos, George. 1998. *Logic, logic, and logic*. Cambridge, MA: Harvard University Press.

Castañeda, Hector-Neri. 1980. The theory of questions, epistemic powers, and the indexical theory of knowledge. In Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein (eds.) *Midwest studies in philosophy* V. Minneapolis: University of Minnosota Press.

Chomsky, Noam. 1972. *Language and mind*. New York: Harcourt Brace, Jovanovich.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1-22.

Chomsky, Noam. 2012. *The science of language: interviews with James McGilvray*. Cambridge, England: Cambridge University Press.

Cowling, Sam. 2017. *Abstract entities*. New York/Florence, Kentucky: Routledge.

Dixon, Robert. M. W. 1963. *Linguistic science and logic*. The Hague: Mouton.

Dik, Simon. 1968. *Coordination*. Amsterdam : North-Holland.

Field, Hartry. 1989a. Realism and anti-realism about mathematics. In Field 1989b.

Field, Hartry. 1989b. *Realism, mathematics, and modality*. Oxford: Basil Blackwell.

Forbes, Graeme. 1994. *Modern logic*. New York; Oxford University Presss.

Garvin, Paul L. 1954. Review of Louis Hjelmslev, *Prolegomena to a theory of language, Language* 30: 69-96.

Gawron, Jean Mark. 1996. Quantification, quantificational domains, and dynamic logic. In Shalom Lappin (ed.) *The handook of contemporary semantic theory* 247-267. Oxford: Blackwell.

Harris, Zelig S. 1968. *Mathematical structures of language*. Interscience Tracts in Pure and Applied Mathematics, 21, New York: Intercience Publishers.

Hjelmslev, Louis. 1953. *Prolegomena to a theory of language*, Indiana University Publications in Anthropology and Linguistics, Memoir 7 of the International Journal of American Linguistics. Supplement to Vol. 19, No. 1. Baltimore: Indiana University Press.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. Preliminaries. In Rodney Huddleston and Geoffrey K. Pullum. (eds.) *The Cambridge grammar of the English language*. Cambridge, England: Cambridge University Press.

Johnson, David E. and Paul M. Postal. 1980. *Arc-pair grammar*. Princeton, New Jersey: Princeton University Press.

Katz, Jerrold J. 1981. *Language and other abstract objects*. Totowa, New Jersey: Rowman and Littlefield.

Katz, Jerrold J. 1984. An outline of platonist grammar. In Thomas G. Bever, John M. Carroll & Lance A. Miller (eds.), *Talking minds: The study of language in cognitive science*, 17–48. Cambridge, MA: MIT Press.

Katz, Jerrold J. 1996. The unfinished Chomskyan revolution. *Mind and Language* 11:270-294.

Katz, Jerrold J. and Paul M. Postal. 1991. Realism vs. conceptualism in linguistics. *Linguistics and Philosophy* 14: 515-554.

Langendoen, D. Terence and Paul M. Postal. 1984. *The vastness of natural languages*. Oxford: Basil Blackwell.

Langendoen, D. Terence and Paul M. Postal. 1985. Sets and sentences. In Jerrold J. Katz (ed.) *The philosophy of linguistics*. Oxford: Oxford University Press.

Martinon, Phillipe. 1926. *Comment on parle en français*. Paris: Librairie Larousse.

May. Robert. 1985. *Logical form*. Cambridge, Massachusetts: The MIT Press.

James D. McCawley, James D. 1981. Everything that linguists have always wanted to *know about logic but were afraid to ask*. Chicago: The University of Chicago Press.

Muller, Claude. 1990. *La négation en français*. Geneva: Droz.

Partee, Barbara H., Alice ter Meulen and Robert E. Wall. 1993. *Mathematical methods in linguistics*. Dordrecht: Kluwer.

Peters, Stanley and Dag Westerståhl. 2006. *Quantifiers in language and logic*. Oxford: Clarendon Press.

Postal, Paul M. 1968. *Aspects of phonological theory*. New York: Harper and Row.

Postal, Paul M. 2003. Remarks on the foundations of linguistics. *The Philosophical Forum* 34: 233-251.

Postal, Paul M. 2004. Junk ethics 2: The most irresponsible passage. In Paul M. Postal. *Skeptical linguistic essays*. 296-305. New York: Oxford University Press.

Postal, Paul M. 2009. The incoherence of Chomsky's 'biolinguistic' ontology. *Biolinguistics* 3.1: 247-266.

Postal, Paul M. 2012. Chomsky's foundational admission. LINGBUZZ/001569.

Potter, Michael. 2004. Set theory and its philosophy. Oxford: Oxford University Press.

Pullum, Geoffrey K. 2007. The evolution of model-theoretic frameworks in linguistics. In James Rogers and Stephan Kepser (eds.), *Model-Theoretic Syntax at 10* (proceedings of the MTS@10 workshop, August 13-17, organized as part of ESSLLI 2007, the European Summer School on Logic, Language and Information), 1-10. Dublin, Ireland: Trinity College Dublin.

Pullum, Geoffrey K. 2011. On the mathematics of *Syntactic Structures*. *Journal of Logic, Language and Information* 20, 277-296.

Pullum, Geoffrey K. 2013. The central question in comparative syntactic metatheory. *Mind & Language* 28: 492–521.

Pullum, Geoffrey K. 2015. What to be realist about in linguistic science. Presentation at the Workshop on Foundations of Linguistics: Languages as Abstract Objects, Technische Universitat Carolo-Wilhelmina zu Braunschweig, Germany, June 25-28, 2015.

Pullum, Geoffrey K. and Barbara C. Scholz. 1997. Theoretical linguistics and the ontology of linguistic structure. *SKY 1997: Suomen kielitieteelisen yhdistyksen vuosikirja 1997* [*1997 Year-book of the Linguistic Association of Finland*], 25–47. Turku, Finland.

Pullum, Geoffrey K. and Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groote, Glyn Morrill, and Christian Retoré (eds.) Logical Aspects of Computational Linguistics:

4th International Conference (Lecture Notes in Artificial Intelligence, 2099) 17-43, Berlin: Springer Verlag.

Pullum, Geoffrey K. and Barbara C. Scholz. 2003. Model-Theoeretic Syntax: Linguistic Aspects. Taught at European Summer School in Logic, Language and Information (ESSLLI), Technical University of Vienna, 2003;

Pullum, Geoffrey K. and Barbara C. Scholz. 2005. Contrasting applications of logic in natural language syntactic description. In Peter Hájek, Luis Valdás-Villanueva, and Dad Westerståhl (eds.), 481-503. Logic, methodology and philosophy of science: Proceedings of the twelfth international congress. London: College Publications Department of Computer Science, KCL.

Pullum, Geoffrey K. and Barbara C. Scholz. 2006. Irrational nativist exuberance. In Robert Stainton (ed.) *Contemporary debates in cognitive science*. 59-80. Oxford: Basil Blackwell.

Pullum, Geoffrey K. and Barbara C. Scholz. 2010. Recursion and the infinitude claim. In Harry van der Hulst (ed.) 113-138, *Recursion in human language*. Berlin: Mouton de Gruyter.

Pullum, Geoffrey K. and Hans-Jörg Tiede. 2010. Inessential features and expressive power of descriptive metalanguages. In Anna Kibort and Greville Corbett (eds.) *Features: perspectives on a key notion in linguistics,* 272-292. Oxford: Oxford University Press.

Riegel, Martin, Pellat, Jean-Christophe and René Rioul. 1994. *Grammaire méthodique du français*. Paris: Presses Universitaires de France.

Rosen, Gideon. Abstract objects. *The Stanford encyclopedia of philosophy*. (Winter 2017 Edition) Edward N. Zalta (ed.)

https://plato.stanford.edu/archives/win2017/entries/abstract-objects/.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, Massachusetts: Doctoral Dissertation, MIT. Published as *Infinite syntax*. 1986. Norwood, New Jersey: Ablex Publishing Corporation.

Rucker, Rudy. 1995. *Infinity and the mind.* Princeton: Princeton University Press.

Sainsbury, Mark. 2009. *Fiction and fictionalism*. New York: Routledge.

de Saussure, Ferdinand. 1959. *A course in general linguistics*. New York: Philosophical Library.

Schlenker, Philippe. 2017. Indexicals. LINGBUZZ/002569. To appear in Sven Ove Hansson and Vincent F. Hendricks (eds.) *Handbook of formal philosophy*.

Scholz, Barbara, Francis Jeffry Pelletier and Geoffrey K. Pullum Philosophy of linguistics. 2011. In the Stanford encyclopedia of philosophy (http://plato.stanford.edu/)

Seuren, Pieter A. M. 1985. *Discourse semantics*. Oxford: Basil Blackwell.

Skyrms, Brian. The explication of '*X* knows that *p*'. *Journal of Philosophy* 64: 373–389.

Stalin, Joseph V. 1950. Marxism and problems of linguistics. Moscow: Foreign Languages Publishing House. (First published in the June 20, July 4, and August 2, 1950 issues of *Pravda*)

Thomasson, Amie, 1999, *Fiction and metaphysics*. Cambridge: Cambridge University Press.

Wetzel, Linda. 2009. *Types and tokens*. Cambridge, Massachusetts: the MIT Press.