Truth, Paradox, and Ineffable Propositions

James R. Shaw
Draft, please don't cite

There is a natural and quite reasonable assumption to make about what it is possible to express in language—roughly: everything.* In this paper I want to present some reasons for thinking this assumption might be false. I'll argue, in a sense to be explained, that there are some things that we can believe which are systematically resistant to linguistic expression in some very peculiar ways.

My motivation for countenancing strong expressive limitations comes from the debates surrounding the semantic paradoxes, where expressive power is a recurring and vexing theme. Each of our many formal theories of truth is under suspicion of being *unable* to express some intelligible semantic concepts or facts—even those dialetheist views that admit true contradictions in the hopes of securing suitably broad expressive power. Despite this, virtually every truth-theorist holds fast to the idea, vaguely understood, that the correct theory of truth will be immune to all worries about expressive power.

I have three goals in this paper. First, I want to help clarify the poorly understood area of linguistic expressive power by giving a clear classification of different expressive limitations, ranked in terms of severity, that are especially relevant to the paradoxes. Second, with the help of a special variant of the Berry Paradox, I want to show just how weak our assumptions about truth need to be to generate some of the very strongest forms of limitation in that classification. Finally, I want to show why granting the existence of these expressive limitations is coherent, and to present some tools we can use to understand and gradually come to accept them. My hope is that these three points will help clear up precisely what trade-offs are required to secure the kinds of expressive power truth theorists typically want and, in the process, to show why admitting the existence of strong expressive limitations is quite possibly the lesser of two evils.

1 A Taxonomy of Ineffabilities

I'll be working with an idealizing assumption and some special terminology. The idealizing assumption is that languages are immutable, isolatable, and well-defined entities in a way that natural languages like English are arguably not.

*In writing this paper, I've benefited from helpful discussions with Sharon Berry, Warren Goldfarb, Peter Koellner, Jon Litland, Bernhard Nickel, an anonymous referee, and from comments from audiences at the Harvard/MIT Eminees Reading Group, the University of Texas at Austin and the University of Pittsburgh.

In particular, I'll take languages to behave roughly like regimented formal languages with fixed arrays of sentences built from a finite lexicon.

In addition to this idealizing assumption I'll be using the terms "concept" and "proposition" in special senses. I'll be calling something a *concept* if it is the sort of thing that stands as the meaning of any kind of sentence part as uttered in context, whatever those might be.[1] Not only nouns and verbs, but quantifiers, coordinating constructions, and prepositions can express concepts on this terminology. I won't dwell on the difficult question as to how to individuate concepts, since this won't affect my arguments.

I'll correspondingly be talking of *propositions* as the meanings of utterances of whole sentences in context, with one provision: that these entities be truth-evaluable or at least candidates for truth-evaluability. Largely for expositional purposes, I will be allowing that propositions may be individuated in ways more finely than by just their truth-conditions or the possible worlds where they are true.[2]

With this terminology in place, I define something to be ineffable in a particular fixed, interpreted language $\mathcal{L}$ as follows.

> *A concept or a proposition is ineffable in* $\mathcal{L}$ if there is no expression of $\mathcal{L}$ which expresses that concept or proposition.

Naturally, I am not offering this definition as a way of clarifying the expression relation—the relation which holds, say, between a given word and the concept it expresses. Rather, I am taking that notion for granted in giving the definition, and grant that in appealing to this intuitive notion my definition may inherit some unclear applications. This is acceptable for my purposes as long as we can pick out some clear cases of expressive limitation. We can do this, and once we do we will see there is nothing necessarily unusual about the existence of ineffabilities in my sense.

To see this, consider the case of Sam and Walter. Sam is from Maine and has never been to Texas. Walter is a native born Texan who has never traveled. Let's suppose these two characters have never interacted with each other, even indirectly, and have never heard of each other. Because of this Sam's idiolect fails to contain anything recognizable as a name for Walter. Arguably then, Sam's idiolect exhibits a conceptual ineffability in my sense concerning whatever it is that a name for Walter expresses.[3]

[1] There are of course different, sometimes overlapping, conceptions of what the meaning of an expression comes to. I am happy to accommodate most of these as candidate specifications of my sense of "concept".

[2] I discuss exactly how the conclusions of my arguments should be adjusted if propositions are coarsened in n.17.

[3] Hofweber (2006) has made an interesting case that if we consider expressibility relative to a broad enough range of contexts, most languages will contain tools like demonstratives which may skirt the simplest kinds of alleged expressive limitations I'm pointing to. For now we can avoid these worries by talking about ineffabilities relative to a language-context pair, or by simply abstracting away from the expressive power afforded by our most contextually flexible linguistic tools, like demonstratives. For a more detailed discussion of how the sensitivities of indexicals may affect my conclusions, see §3.2.

If we concede this, we will recognize that there are vast arrays of conceptual and propositional ineffabilities in not only Sam's idiolect, but in our own. Many of these ineffabilities will, like Sam's missing name for Walter, be relatively innocuous. But others might not be. To bring out when an ineffability is of interest, I'd like to draw three distinctions, each of which tracks one dimension along which we can assess an ineffability's strength.

First, we should note a difference between the expressive gaps connected with concepts and propositions.

> *A conceptual ineffability in* $\mathcal{L}$ is the ineffability of a concept in $\mathcal{L}$.

> *A propositional ineffability in* $\mathcal{L}$ is the ineffability of a proposition in $\mathcal{L}$.

Conceptual ineffability generally entrains propositional ineffability. To see this, consider Sam's idiolect again. Because Sam cannot directly talk about Walter, he cannot directly ascribe properties to him. He cannot, for example, say that Walter is gaunt, or an epicure, or born in Albuquerque. His conceptual ineffability here generates a series of propositional ineffabilities. And the phenomenon seems general: whenever a set of propositions has their expression facilitated by a given concept, then inexpressibility of the concept can bring the inexpressibility of those propositions along with it. It is not obvious that something like this relation holds in reverse.

The second, and most important distinction I'd like to make, concerns whether an ineffability can be overcome by an expansion of the language. Call any interpreted language $\mathcal{L}'$ an *extension of a language* $\mathcal{L}$ if it has at least the conceptual resources of $\mathcal{L}$: any concept c expressible in $\mathcal{L}$ is also expressible in $\mathcal{L}'$. Then the following definitions track an important sense in which an ineffability can be avoided.

> An ineffability in $\mathcal{L}$ is *removable* if it is absent from an extension of $\mathcal{L}$.

> An ineffability in $\mathcal{L}$ is *essential* if it is present in all extensions of $\mathcal{L}$.

Most kinds of ineffability in a given language which immediately come to mind are removable. Consider, for example, English circa 1500 ACE. This was a language rife with conceptual and propositional ineffabilities concerning the subject matter of chemistry. Nonetheless, each such ineffability was removable in my sense, as is witnessed by present day idiolects of chemists.

Essential ineffabilities, if any exist, are clearly more problematic than removable ineffabilities. But we cannot assess just how problematic without drawing one further distinction among types of essential ineffabilities. We must distinguish cases where a language unavoidably fails to express one particular proposition or concept from cases where a language unavoidably fails to express, collectively, a group of propositions or concepts.

> *A singular essential ineffability in* $\mathcal{L}$ is the ineffability of a particular concept or proposition that is present in all extensions of $\mathcal{L}$.

> *A class-based essential ineffability in* $\mathcal{L}$ is a set of propositions and concepts S such that
>
> (i) Every $s \in S$ is ineffable in $\mathcal{L}$, and
>
> (ii) In every extension $\mathcal{L}'$ of $\mathcal{L}$, some $s \in S$ is ineffable in $\mathcal{L}'$.

Let me give an example to help bring out how this last distinction operates. Consider the set of propositions P which contains for each real number r, the proposition that r is a real number. So, for example, the proposition that π is real, the proposition that 9 is real, and the proposition that $\frac{17}{5}$ is real are all in P. If we assume for now that these are all distinct propositions, then P is a set with cardinality of the continuum. If we suppose, further, that the languages up for consideration are those in-principle learnable by human beings, those languages will have only a countably infinite number of sentences—a smaller number of sentences than P. Consequently if we start with some interpreted language $\mathcal{L}$, there will be an uncountable subset P' of P, such that every proposition in P' is ineffable in $\mathcal{L}$.[4] Now, any propositional ineffability from P' in $\mathcal{L}$ can be removed at once (barring practical difficulties): all we have to do is add a name for the relevant real. But we cannot remove *all* the propositional ineffabilities from P' in $\mathcal{L}$ *at once*. This is because no matter what language $\mathcal{L}'$ we pick as our extension of $\mathcal{L}$, $\mathcal{L}'$ will exhibit the same limitations as the original language—it will have only countably many sentences. Consequently, some propositions in P' must go unexpressed in $\mathcal{L}'$. So P' is an example of what I want to call a *class-based essential ineffability in* $\mathcal{L}$.

Again, the distinction marks a difference in strength. The existence of a singular essential ineffability entails the existence of a class-based essential ineffability—namely, that given by the singleton containing the concept or proposition singularly inexpressible. But the existence of a class-based essential ineffability needn't entail the existence of a singular essential ineffability. The example I just gave based on talk of real numbers should show how the former might arise without the latter. The difference in strength here helps account for the fact that class-based essential ineffabilities have not always been seen as detrimental.[5] Singular essential ineffabilities, however, tend to be strongly resisted.[6]

[4] Again, integrating context complicates things. I suspect indexicals won't ultimately avoid limitations from talk of reals, but even if it does, this shouldn't detract from the utility of the example.

[5] Class-based essential ineffabilities arise very naturally in formal theories of truth which fragment semantic and pragmatic properties. And, in fact, several philosophers have explicitly argued for the existence of class-based essential ineffabilities independently of the paradoxes. For a quick version of a cardinality argument for class-based singular ineffabilities governing properties, see Tye (1982). Camp (2006) pp.14–16 presents another very different set of considerations favoring class-based essential ineffabilities in connection with metaphorical contents.

[6] Searle seems to have thought singular essential ineffabilities in general are not possible, at least on one reading of his 'principle of expressibility' of Searle (1969) §1.5. Schiffer denies the possibility of their being singular essential ineffabilities governing properties at p.71 of Schiffer (2003). The claim that singular essential ineffabilities don't exist is typically assumed rather

The three binary distinctions I have made above create a six-fold classification of ineffabilities and a rough rank ordering in strength.[7]

Increasing Strength ↑	Singular Essential Conceptual Ineffability
	Singular Essential Propositional Ineffability
	Class-based Essential Ineffability (some conceptual ineffabilities)
	Class-based Essential Ineffability (only propositional ineffabilities)
	Removable Conceptual Ineffability
	Removable Propositional Ineffability

2 Paradoxes of Ineffability

In this section I'm going to develop a family of paradoxes that I'll use to argue for the existence of what are 'almost' singular essential propositional ineffabilities—the second strongest form of ineffability on my taxonomy. I'll have to defer a discussion of my caveat 'almost' to §§3.2–3.3. For now, it is safest to think of what I'm doing as arguing directly for singular essential propositional ineffabilities *simpliciter*. Even more controversially, I'm going to try to do this 'constructively', by trying to produce a proposition which can be coherently thought but which cannot coherently be expressed in language. Since this idea is likely to appear extremely radical, I want to motivate the argument, and its assumptions, in several steps. Let me begin by first saying a few words about my methodology.

2.1 Methodology

How could it even be possible to *argue* for the existence of a particular inexpressible proposition? Wouldn't any attempt be self-undermining?

My claim is that there is a proposition which can be thought, but which cannot stand as the conventional, literal linguistic meaning of the utterances best suited to express it. It's obviously hopeless to argue for the existence of such a thought in any conventional way—by deducing how one would state it, for example. This means I will have to rely on a special kind of open-mindedness and cooperation from my readership. I will try to convey my thought in an indirect way: by noting there should be a 'limiting case' of a series of other intelligible, and sometimes conventionally stateable, thoughts. To ascertain whether this process is effective, I must ask my readers only to consider two things: do they

than explicitly argued for in discussions of truth. See, for example, Priest (2006) p.54. Very few philosophers have openly condoned the existence of singular essential ineffabilities. On certain controversial readings, Frege (1951) and Wittgenstein (1961) concede their existence. Lewis (2001) §2.3 more clearly argues for the position. I discuss his case in §3.

[7] For another recent classification of expressive limitation which overlaps to a certain extent with mine, see Kukla (2005). Also, Hofweber (2006) gives a set of important distinctions I gloss over here concerning context.

find they can understand the relevant limiting case of my series of thoughts? And, if so, can they find a conventional way of expressing it in literal language?

But isn't it simply incredible that this could happen—that one could think something without being able to say it? Of course I agree that this seems implausible, but I believe that it happens on extremely rare occasions because of a particular kind of linguistic technicality. To help see what I mean by this, and to understand what I am doing a little better, consider what might seem like a comparably implausible circumstance: that I could pose any person P a question with many true answers that both P and I could easily recognize, but which neither P nor I could state as true answers to the question.

I might convince Smith of this simply by asking the right question: "What is an example of something neither Smith nor I will speak of in giving an answer to this question?" It's easy to see the problem for Smith and me. Being something spoken by either of us in answer to the question immediately disqualifies it as producing one of the question's true answers. But we can easily see there are very many true answers to this question, and even think of some particular examples: I might point at a badger, winking at Smith. ""Badger" is a perfectly correct answer to that question," we might think "but of course one we could not state as a true answer." We could be perfectly right on both counts.[8]

Smith and I are unable to give true answers to the question in speech on a technicality. Nonetheless, we can easily think true answers to the question without a problem. I'm *not* claiming that this case is perfectly analogous to the one I'm about to produce. But it can be helpful to understand what I'm trying to do. I think that a similar, somewhat more complex, linguistic technicality precludes the expression of certain propositions, but that there is no bar to our thinking them. And there is no bar to *convincing* someone of a particular example, even without stating it.

The way I think all this can be accomplished is by modifying and 'strengthening', in some important ways, the well-known Berry Paradox. Before I discuss the strengthenings, let me make some remarks about the Berry paradox itself.

2.2 The Berry Paradox

The immeasurable complexity of the world seems bound to outstrip the immediate resources of any one fixed language. Reflecting on this truism, we can arrive at the conclusion that our language as it stands may easily fail to express everything there is to express. Interestingly, some attempts to state these obvious and weak expressive limitations only succeed if they somehow surpass the very limitations they aim to report. The kind of circularity I'm alluding to here is evinced by the well known Berry paradox, which goes roughly as follows.[9] Imagine all the definitions which pick out an integer using fewer than twenty English words. While this list is finite, there are infinitely many integers to speak of. So there looks to be a set of integers not defined by any sentence on

[8]The example here is inspired by the discussion of 'vagrant concepts' in Rescher (2009).

[9]The origin of the paradox is Russell (1908), who attributes it to the eponymous Bodleian librarian.

the list, and a least integer among them. We can try to pick this integer out by saying "Let n be the least integer not definable in fewer than twenty English words." This definition, if it has succeeded, has defined an integer not definable in English with fewer than twenty words, in fewer than twenty English words. So something has gone wrong.

What *exactly* has gone wrong is a matter of dispute. In general, resolutions of the Berry paradox parallel resolutions to the much more discussed liar paradox. I don't want to survey the array of options available here. Instead, for heuristic purposes, I want to show how our response to the Berry paradox is radically constrained on a particular set of naive assumptions:

(B_1) There is a speech act of definition which is not language dependent.

(B_2) There is a set of actual attempted definitions of integers which (simply) succeed. If a definition is successful in this way, there is a unique integer it defines.

The idea motivating (B_1) is simple: there is a speech act through which we associate, by stipulation, certain signs with entities the speech act somehow picks out. This can happen in English, French, or Chinese—the speech act of definition is not tied to a particular language. Focusing on definitions of integers, (B_2) expands on this idea by noting that the speech act of definition, like other speech acts, can succeed or fail. I might *try* to define an integer when things get in the way. This might occur, for example, if I said "Let n be the integer Eleanor just mentioned" when Eleanor mentioned no integer whatsoever. But (B_2) makes a further claim: there is some set of attempted definitions of integers which, simply, succeed in defining an integer.[10] For example, if we considered the set of all utterances in any language which are possibly construed as attempted definitions, that set of utterances can be divided into those definitions which succeed in picking out a unique integer and those which do not.

Though these assumptions naively seem unproblematic, they are controversial because of the paradoxes. My goal here is not to defend these particular assumptions, but to trace out their consequences for heuristic purposes. One of their consequences, drawing on the Berry paradox, is this: that there are integers bearing certain properties that sometimes cannot be picked out with our standard way of talking about those properties.

This is likely to sound confusing, so let's begin by looking at a relatively simple case. Consider two characters, Mika and Ahmed. Ahmed is, for whatever reason, in the business of producing definitions of integers today. Mika, a defender of principles (B_1) and (B_2), informs Ahmed that on her view all his attempted definitions will either succeed or fail, and if any succeed there

[10]The caveat "simply" is to allow for broadly dialetheist positions on which a definition may both succeed and fail. The definitions which simply succeed are not among those. Also, it is worth noting that this assumption may look unattractive when we consider issues of vagueness. But for the purposes of this paper, we can provisionally abstract away from such considerations. Vagueness may at the end of the day actually assist the case for most kinds of ineffability.

will only be a finite number that do so. Consequently there will be an integer one greater than the greatest integer he defines—an integer he won't succeed in defining that day.

In normal circumstances, this conclusion will be completely appropriate. Suppose the only utterances remotely resembling attempted definitions which Ahmed produces are the following two.

(i) "Let n be the smallest positive integer which can be written as the sum of two cubes in two different ways." [n=1729]

(ii) "Let n be Douglas' favorite integer" [n=42]

Then the number Mika is after is 1730. This is so even if Ahmed were to produce some other definitions which clearly fail. For example if he were to have also said only:

(iii) "Let n be Mark's son's favorite integer." [–]

when Mark has no son. 1730 is again the number Mika is after.

But is this always the case? Suppose Ahmed grasps Mika's point, but triumphantly announces that he can circumvent her conclusion. He boldly announces his final attempted definition of the day:

(iv) "Let n be the integer one greater than the greatest integer I successfully define today." [n=?]

If n defines any number, it would seem to be one greater than itself. There are no such integers. So what should we conclude? Some might conclude that we should reject (B_1) or (B_2) and the reasoning that led to Mika's claim. But this is not *obligatory*. Mika may stick to her guns in a quite natural way. We can reason to the conclusion that there is an integer n greater than itself only on the presumption that Ahmed's (iv) involves a successful definition. So, to retain (B_1) and (B_2), we need only deny that his utterance is among the simply successful definitions of the day. In fact, *even if* we reject (B_1) or (B_2) we are probably going to end up endorsing something like that claim *anyway* (for example, by denying the coherence of the concepts Ahmed is trying to employ in his definition).

If we merely treat Ahmed's last utterance as a failed definition, we end up again with a case which clearly vindicates Mika's thought that there is an integer one greater than the greatest Ahmed successfully defines that day: it remains 1730.[11] We have four attempted definitions, two clearly succeeding, two clearly failing.

If we maintain (B_1) and (B_2) and reject the success of Ahmed's utterance of (iv), there is one obvious cost. It seems that if there were an integer of the

[11] If we reject the success of (iv) by rejecting (B_2) we may only end up rejecting the *reasoning* that led to the truth of Mika's claim, not the claim itself. If so, her claim seems clearly true. We might also end up rejecting the coherence of Mika's final claim if we reject (B_1)—but this is difficult to maintain when we *see* how the results of applying our view generate a unique integer of the kind Mika intuitively desired.

kind Mika alludes to, Ahmed's utterance *should* successfully define it. Ahmed's utterance looks tailor-made to pick out the relevant integer if there is one. This is the radical conclusion that I mentioned was a consequence of (B_1) and (B_2): there are integers bearing certain properties that sometimes cannot be picked out with our standard way of talking about those properties. One might worry whether this position can be elaborated coherently. In particular, one might wonder whether there can be consistent formal accounts which permit such special kinds of descriptive failure. This worry is reasonable, but answerable. In fact, there are several such accounts to choose from. I have in mind context sensitive accounts like Glanzberg (2001, 2004), and token-sensitive accounts like Gaifman (1992).[12] Each of these theories provides different, systematic ways of explaining why some utterances containing semantic terms fail to express what they seem apt to when paradox arises. Glanzberg accomplishes this by positing a special interaction between contexts of utterance and (sometimes tacit) quantifier domains. Gaifman, by contrast, locates the phenomenon in the operation of semantic words themselves. Both theories have the virtue of explaining the temptation to think the relevant defective utterances 'should' express something coherent: because distinct utterances of the very same sentence types sometimes *do* function normally, just as long as they avoid degenerate self-reference.

In fact, the case I've been examining is very much like those paradigmatically used to motivate such views. To see this consider what would happen if *Mika* were to utter (iv) the day in question, or if Ahmed were to utter it the day after (using "yesterday" instead of "today"). There would be nothing paradoxical about such utterances: supposing them to define an integer n (in this case 1730) does not lead to contradiction of itself. The utterances doing the defining are no longer among those being used to generate the integer, hence there is no reason to suppose that the definition requires an n greater than itself, or is in other ways paradoxical. Similar remarks apply to the *claims*, and not the definitions, made by anyone including Ahmed, that there is a unique integer of the relevant sort. Likewise, there is nothing yet paradoxical about the remarks I have been making in this section, on Mika's behalf. This helps motivate the idea that what is bad about Ahmed's attempted definition is not necessarily his choice of words, or his acquiescence in the language-independent concept of definition Mika endorses. It is the relationship between Ahmed's utterance and the properties he wants to use to characterize an integer.

Incidentally, it is worth mentioning how difficult it is to reject (B_1) or (B_2) to block Mika's conclusion in the surveyable case I've just given. It is hard to deny that there is *some* notion of definition, call it definition*, on which it is appropriate to say that Ahmed's (i) and (ii) succeed in defining* integers whereas his (iii) does not. *Whatever* that notion is, it seems like it will either be one on which Ahmed's (iv) succeeds or fails as a definition*. If the utterance succeeds (say by fragmenting the notion of definition), it should pick out a unique integer (say 1730) on pain of being unrecognizable as a notion of definition at all. If

[12]Whether Glanzberg would be happy to apply his apparatus to this case in the way I'm suggesting is not obvious. I simply want to note that the system may exhibit the kind of compositional flexibility that maintaining (B_1) and (B_2) requires.

so, it opens up an avenue for the claim Mika is after to be intelligible and to hold, in some intuitive sense (e.g., she is talking about the number 1731). If, on the other hand, we jettison Ahmed's utterance as a successful definition*, Mika's utterance again seems, in some intuitive sense, after a fact about the relevant speech acts made true by the number 1730. Thus, either way, we end up restructuring the case in ways that makes Mika's motivations and her claim look entirely appropriate. If *all* cases can be thought of in these ways, it makes it look like there is an appropriate *general* truth about the speech act of definition (involving definition*, and other notions if there are any) justified by the kind of reasoning Mika goes through.

In any event, I hope to have established my intermediate heuristic conclusion: maintaining (B_1) and (B_2) requires the existence of integers that bear certain properties even though, sometimes at least, we cannot pick out those integers with words normally picking out those properties. If we extend these considerations to the Berry paradox itself, we will endorse the conclusion that there is a least integer not successfully definable in fewer than 20 English syllables, but will reject the claim that the Berry definition successfully picks that integer out.[13]

It is very important to note that *even if* we grant (B_1) and (B_2), we are a long way from having any interesting ineffabilities of any kind whatsoever. First, definitions are not claims. They don't express propositions, so there is no propositional ineffability in the offing. Moreover, barring certain practical difficulties, there is no *particular* integer that Ahmed cannot define, hence no singular essential conceptual ineffabilities either. In fact, again barring practical difficulties, it doesn't even seem like there is a *class-based* essential ineffability here. Ahmed has enough linguistic resources to recursively name *all* the integers.

So to repeat: even on the assumptions I've given, the Berry Paradox does not obviously generate any interesting worries about expressive limitation. What I want to show now is that we can twist and modify the formulation of the Berry Paradox in such a way that strong ineffabilities do threaten to appear.

2.3 Relative Ineffability Paradoxes

Let me start by describing the two key assumptions of my argument, that are helpfully thought of as analogs of (B_1) and (B_2).

(I_1) There is a notion of truth which is not language dependent.

(I_2) There is a set of actual utterances produced in assertion which are (simply) true.

[13] As before, this might be a little overly simplistic. If we think of definitions as speech acts, those acts might easily be context sensitive (like "let n be the number Eleanor just mentioned"), in which case even one English expression could define several integers relative to different contexts. For these reasons the Berry argument, and its conclusion, would have to be adjusted in ways I've alluded to earlier.

The motivations for (I_1) are, as before, fairly straightforward: There is a speech act we call "assertion" which we typically use to do something like convey information. When we produce an utterance in assertion, and things stand favorably in some way (perhaps when things are the way things are asserted to be, or when we have performed our action according to appropriate norms), our utterances are appropriately spoken of as being true. There are not multiple different speech acts, and multiple different success conditions, for different languages. Just one thing, assertion, takes place in English, French, or Chinese. Just one virtue is possessed by these assertions when what is asserted is done appropriately, or in favorable circumstances. Accordingly, there is a single notion of truth which applies to a single kind of speech act—assertion—that can take place in multiple different languages.

(I_2) requires a little more clarification. First, to endorse (I_2) isn't to take a stand on the issue of whether utterances are the 'primary' bearers of truth. Nor does it require denying that other entities can be truth-bearers—thoughts, for example. (I_2) only relies on the idea that it can make sense to talk of utterances used in assertions as being true or not, even if what makes this the case is a relation between the utterances and some more fundamental bearer of truth.

Second, to endorse (I_2) is not to presuppose that truth is a *property* in any substantive sense. It is entirely compatible with views on which there is 'no one thing which makes every true utterance true', as some deflationists might hold. (I_2) only claims that, in any imaginable circumstance, there will a set of utterances produced in assertion which are true, regardless of what makes that the case. If even this sounds too committal, (I_2) can be read as a claim about appropriate use: in any imaginable circumstance there is a unique set of utterances such that it would be correct or appropriate to ascribe the word "true" to them. In other words, (I_2) simply claims there are coherent, exhaustive norms for the correct use of the word "true", at least as concerns utterances produced in assertion, and that those norms (if only indirectly, through use of the word) privilege a certain set of utterances.

Third, for all (I_2) claims, there may also be utterances which are neither true nor false. There may even be utterances which are both true and false. These are not in the set of which (I_2) speaks. What I mean by adding the rider "simply" to true utterances is to pick out those which it is appropriate or correct to speak of using the word "true", but not appropriate or correct to speak of using words like "not true" or "false". Again, supposing this makes sense is simply to suppose that there are coherent, exhaustive norms for the use of the word.

Fourth, if you are tempted to deny (I_2) because you think that there aren't *yet* coherent norms governing the use of the word "true", construe (I_2) as being about our best possible theory of truth (or any one of our best possible theories if we have choices) and not our actual, allegedly defective or incomplete, practices.

Finally, if you think that "true", or typical utterances containing it, are context sensitive, construe (I_2) as a claim to the effect that the norms governing the use of "true" again privilege a set of utterances—this time those that in *some* context it can be appropriate or correct to speak of as (simply) true.

Just as before, I can imagine that both (I_1) and (I_2) could seem controversial because of the paradoxes. Though I won't be able to do full justice to the claim here, I think rejecting either of these theses is potentially devastating for work in the philosophy of language, philosophical logic, and of course work on truth itself. It should be obvious why rejecting (I_2) has this effect. It threatens to make it pointless to study truth. If there is no possible coherent and complete set of norms for how to talk about assertions using the word "true"—even norms which tell us to 'ignore', 'remain silent', or 'say what we will' with respect to problematic utterances—what other conclusion could we draw than that truth is an incoherent notion not worth anything like the kinds of serious, detailed formal study it has long received?

Denying (I_1) is not quite as devastating, but in addition to seeming implausible, it threatens to demolish a vast array of important projects in the philosophy of language and logic. Admitting the existence of language-independent notions like truth or inference, and the ability to talk about them, are presupposed by foundational accounts of meaning and intentionality. The intelligibility and coherence of these foundational accounts, if they are to be general, requires just such universal properties. Jettisoning the language-independent notions seems to force accounts of meaning and communication to be incomplete in ways that will leave intentional practices of human beings without any *possible* explanation.

For reasons like this, I find (I_1) and (I_2) inescapable. If you do not feel similarly compelled, though, my argument to follow can be reworked as a conditional one, tracing out what follows from these assumptions. The resulting conditional conclusion will still be an important one, since it will constrain trade-offs between the structure and breadth of our theories of truth and expressive power in ways that may be illuminating, even for those tempted to deny its antecedent.

Let's at least provisionally draw on these assumptions, then, to construct a puzzle with the help of Ahmed and Mika. Ahmed has reflected on Mika's understanding of definitions and wants to similarly draw out some consequences from his commitment to a notion of truth governed by (I_1) and (I_2). He reasons as follows of Mika: she will produce at most finitely many utterances today. We can expect some of these utterances might be among those which are simply true. Among those utterances some may make an existential claim. And among *those* utterances some may be made true by unique integers.[14] We can grant that this set of simply true existential claims made by Mika today, each of which is satisfied by a unique integer, might be empty. But we can say with certainty it will be finite. Because of this, if we consider the integers satisfying the existential claims in the set collectively, there will be a greatest one—call it n. But then there will be an integer, $n' = n + 1$: a number one greater than the greatest number uniquely satisfying a simply true existential claim of Mika's.

Let's reformulate and elaborate on the reasoning a little with the help of some definitions.

[14]This is in effect supposing that a language-independent notion of satisfaction exists which parallels that for truth. If this seems somehow more substantive than (I_2), one can add it as a premise for the reasoning to follow.

> The *e-referent of an utterance* U is n if U is a simply true existential claim uniquely satisfied by n, and undefined otherwise.

For example, suppose Mika says "There is a prime number between 3 and 7," then the e-referent of her utterance is 5. If Mika, at another time says "There is a prime number between 1 and 10," then although her utterance is true it has no e-referent. This is because there are too many integers satisfying her claim and so no particular integer is singled out by it.

Here's a second definition.

> The *super referent of a set of utterances* $\mathcal{U}$ is n if n is one greater than the greatest integer e-referred to by an expression in $\mathcal{U}$, and 7 if there is no such integer.

In effect, the super referent of a set $\mathcal{U}$ 'diagonalizes out' of the set of e-referents in $\mathcal{U}$. Take a quick example. Suppose Mika makes the following claims in succession:

(a) There's a number between 1 and 3.

(b) There's a number between 4 and 6.

(c) There's a number between 7 and 10.

Letting $\mathcal{U}$ be the set consisting of those three utterances, we can see the super referent of $\mathcal{U}$ is 6. This is because only the first two utterances in the set e-refer, and their e-referents are 2 and 5 respectively. So 5 is the greatest integer e-referred to in the set, and 6 is its super-referent. There may fail to be a greatest integer e-referred to for at least two reasons. Perhaps no utterances in a set $\mathcal{U}$ e-refer. Or perhaps the set contains utterances which e-refer to arbitrarily large integers. Either way, the super referent of such a set is 7.

Back to Ahmed and Mika. Recall, Ahmed endorses (I_1)–(I_2). These, along with some basic facts about existential claims, led him to a conclusion which might have been put by saying there is a unique integer which is the super-referent of Mika's utterances today. But as in our discussion of the Berry paradox we can doubt this. Suppose that after (a)–(c) Mika utters only (d).

(d) There's a unique integer which is the super-referent of Mika's utterances today.

(d) seems to restate the conclusion that Ahmed came to. But it can't be simply true. If it were, it would clearly be satisfied by a unique integer n. But, given what (d) says, n should then be the super referent of the set of utterances consisting of (a)–(d). But such an n would be greater than any integer uniquely satisfying an existentially quantified utterance of Mika that day, n included—contradiction.

What should we do? One option is to reject one of (I_1) and (I_2). As I have said, I believe this would be utterly disastrous for the project of giving a theory of truth. Fortunately as before, we have another option: reject the claim that (d) is simply true. Maybe it's false. Maybe it's neither true nor false. Maybe it's true *and* false. Maybe it has some other more complex status. Maybe we simply must 'be silent' concerning it. As long as it isn't simply true—as long as it isn't among the utterances we should be speaking of as "true" and nothing but—it isn't in the set of utterances relevant to the determination of super-reference. If this is the case, the threat to (I_1) and (I_2) drops away.

As before, there is one more striking cost. It looks like, once we consider (d) less than perfectly and simply true, the reasoning Ahmed goes through is justified and there is a unique super referent of the four utterances of (a)–(d) after all—6. But isn't that what (d) *says*? In a manner analogous to the Berry paradox, if we want to maintain our principles (I_1) and (I_2), we must accept this: though Mika's utterance of (d) *looks* like it of the right form to express the fact that is the conclusion of Ahmed's reasoning, it cannot. And just as before, there are potential asymmetries in accounting for different uses of the very words which are, somehow, defective in Mika's mouth. If Ahmed utters (d), there is nothing that need be problematic about his utterance. Such an utterance can in principle, and non-paradoxically, be counted as an uncontroversial, pure truth, as several theories of truth would have us do. Likewise for the utterances I've been making so far. Again, this is because Ahmed's utterances and mine are not among those discussed by (d). This again may foster the sense that whatever imperfection Mika's utterance exhibits is due to the relationship between the proposition attempting to be stated, and the methods for stating it.[15]

The case of Mika and Ahmed thus reveals a parallel consequence to the one reached in §2.2: there are certain propositions which one cannot conventionally express with the words we would take to best express them. If we accept (I_1) and (I_2) there is a basic fact about the existence of a super-referent of Mika's utterances on a particular day, something simply true—not neither true nor false, not both true and false, not half-true—that holds regardless of what Mika's utterances are like. Mika's (d) looks like it should express the proposition whose truth turns on the fact in question. But Mika's utterance cannot be simply true. So Mika's utterance doesn't express the relevant proposition.

We may allow that individuals other than Mika may express this proposition, or that Mika may express it at other times, drawing on potential asymmetries between various utterances of sentences like (d). And as before, we have formal models which reveal how this is possible, if we choose to avail ourselves of them. This would allow us to view the expressive limitation faced by Mika, however we classify it, to be quite limited in that it would be confined to a particular person on a particular day only. This is why the paradox witnessed here might be called a *relative* one. Shortly, I'd like to show that further elaborations of the paradoxes may not be so easily quarantined.

Before we turn to that task, though, it is important to pause and fully take in the details of this limited case. We must ask ourselves: *given* that

[15]These asymmetries may disappear, of course, if Mika begins to make claims about Ahmed's utterances, or my own, as we will see. This doesn't change the intuitive correctness of Ahmed's, potentially unspoken, conclusion *given* our assumptions.

there are coherent norms governing the use of the word "true", which of Mika's utterances of (a)–(d) are among those to which it is appropriate or correct to apply that word? Is there any reasonable way of using of "true" on which it is appropriate, in any context, to speak of Mika's (d) as "true" (without it also being appropriate to speak of it as "untrue", "false", etc.)? I find it unbelievable this should be the case, given that there is no single integer we could possibly see as satisfying the existential claim. But even if one feels that one such integer *can* be picked out by the claim, this is fine. The point is that we can either group this use of (d) in with the existential claims simply true on account of unique satisfaction by an integer, or not. Either way, we can 'diagonalize out' of the integers associated with Mika's utterances, however many of them e-refer. There is a *kind* of integer which arises in this case from a particular procedure. The fact that it arises, granting (I_1) and (I_2), seems inevitable. That is *just* what it means to claim there is a unique super referent of Mika's utterances. It is important for what follows to genuinely ask yourself (again, *given* that there are coherent norms governing the use of "true") is there a regular procedure that generates some integer in this way, and if there is, isn't there a *fact* here, and a *proposition* which tracks whether that fact obtains or not. Can you think the thought "*this* kind of procedure generates a unique integer in this case"? Again, granted coherent norms for the of use of "true", I find the procedure intelligible, and its outcome in Mika's case inevitable. I have to belabor this point because, as I alluded to before, I am trying to 'encroach' on an ineffable truth—one which by my own admission I will not be able to state. The only way to get you to think this proposition is by alluding to it, as a kind of limiting case of propositions thinkable (and, on my view, sometimes stateable) in other cases like this one. Let's now turn to this final step.

2.4 Absolute Ineffability Paradoxes

We've now seen how one particular set of utterances made on a particular day generated a unique integer which was their super-referent. But there was obviously nothing special about that one set. We could have picked out any set of utterances to generate other super-referents. For example, we could have picked the utterances made by someone else on a different day, or those made by anyone over three feet tall while standing next to a lamppost in Hartford in the 21$^{\text{st}}$ century. As before, we would expect the utterances which are generating the super-referent to be poorly placed to make claims as to the existence of that very super-referent. This was one of the consequences of accepting (I_1) and (I_2). This suggests a natural way of pushing the paradox to its counter-intuitive limits: by considering the set $\mathcal{U}$ of all utterances ever actually made, by anyone, in any language whatsoever. Then we can consider the following paradoxical claim which, though not simply true by my own admission, will be instrumentally useful in getting at the proposition I find intelligible.

(A) There is a unique integer which is the super-referent of $\mathcal{U}$—the set of all utterances ever made.

If (I_1) and (I_2) obtain, it is easy to see why one might be inclined to utter this sentence. From among the utterances ever made, some will be among the plain old truths of existential form, satisfied by unique integers. The set of integers picked out in this way will either have a greatest integer or not. Either way we have a super-referent for the set of all utterances. So there *is* a unique integer which is a super-referent of $\mathcal{U}$.

Of course things aren't quite so simple. What I just said can't be straightforwardly true. Grant me for the moment that the set of utterances over all time won't e-refer to arbitrarily large integers.[16] If my utterance of (A) was just plain old true, it would be satisfied by the super-referent of all utterances ever, including the one I just produced. Whatever that integer is, it will be greater than itself—and we can't have that. So I will concede: what I have said isn't straightforwardly true. In fact, *no one* could utter (A), at any time, in any place, and be speaking a straightforward truth. Nonetheless, in uttering (A) as I do, I hope to be *getting at* something—something which *is* straightforwardly true. (Though I grant my saying *that* might be problematic as well.)

Consider the circumstance of Mika from earlier. She could have been aware of the truth at which Ahmed arrived by his reasoning. She could have gone through that reasoning herself. There was thus a proposition—a true proposition pure and simple—which we and she could *witness*, which Mika would be at pains to express. The best candidates for her to express the proposition in question were utterances of (d)—utterances doomed to falter *for her* in some way. Nonetheless Mika might find it useful to utter (d), with suitable reservations and hedging, if she were trying to get across—in some non-literal way—the facts she was after. I, as someone who endorses (I_1) and (I_2), find myself similarly pained and similarly disposed to produce the problematic utterances.

As before, though the set of all utterances of any language is far from surveyable, we can 'see' and understand the grounds for the truth of the generalized proposition about super-reference. Once we adopt (I_1) and (I_2), it seems that however we flesh out our view about how to use the word "true", the super-referent of the set of all utterances is bound to arise. It is even *in principle computable* just as we located the specific super-referent of Mika's utterances despite her defective pronouncements—it is just a matter of tallying more utterances than before. The would-be result of such a computation might even be a number that someone might successfully refer to, for example, without making an existential claim.

The reason I must acquiesce in producing what are manifestly problematic utterances to try to convey the truth I believe is correlated with (A) is that it is *constructed so as to preclude* literal and truthful expression. What's happening here is in some ways analogous to the case of the 'unstateable true answers' to the question I hypothetically posed Smith in §2.1. The extent to which an utterance *could* qualify as stating the proposition in question is the very extent to which it is disqualified from being able to state it. Why? Because it's clear that the

[16] This seems highly likely to be true. If not, we can redefine e-reference using well-orderings of sets of elements of higher cardinality to achieve the desired problem.

proposition that is the limiting case of that exemplified by Ahmed and Mika's case is an existential one—it's a thought to the effect that something *exists*. Any claim which could qualify as stating that proposition would therefore be an existence claim of some kind. The proposition in question is a simple truth about the existence of a certain kind of integer. But no simply true existential claim can state the proposition in question on pain of generating an integer one greater than itself. So the conditions required for an utterance to state the proposition in question is precisely what is required to show that utterance does not state the proposition in question.[17]

My claim is that if we consider a series of cases like those involving Mika and Ahmed, we will see a 'series' of propositions which are increasingly difficult to express, and a limiting case of these propositions which is inexpressible. There are two simple ways to block this process: to claim something 'changes' in the limit, or to say that there was something problematic about what I concluded in the more basic cases like those of Mika and Ahmed.

The problem with the first of these options is that there is actually nothing very special about the limiting case from a structural perspective. The limiting case here is actually not in any way different in *kind* from the case given in §2.3. The only difference is that the proposition I'm trying to convey concerns more utterances than before. Recall that all the propositions we are considering effectively claim that a process applied to a group of utterances generates an integer. What we'd like to understand is why this process does indeed apply in this way to certain groups of utterances but not others. Just adding more utterances doesn't seem relevant to whether or not the process will succeed, or make sense. So we need some special explanation of some interesting change in the limit. I know of none.

This leaves us the only other option of going back to the case of Mika and Ahmed and redescribing the case to allow for (I_1) and (I_2), but not allow that there is an integer one greater than the greatest integer satisfying a simply true claim of Mika (or perhaps to not allow for the coherence or intelligibility of that claim). The challenge is to give a coherent, complete set of norms for the word "true" applicable to any assertion, such that when we narrow our focus to the simple truths Mika uttered from among (a)–(d), and the integers satisfying the existential claims in that set, there is not an integer one greater than the greatest among them. Or we would have to convince ourselves that even though there is a coherent complete set of norms for the use of "true" the procedure of generating integers I'm alluding to simply doesn't make sense, and is somehow incoherent.

[17] I've been working on the assumption that propositions are structured, but this assumption is not obviously necessary for the importance of my conclusion. For example, a possible worlds theorist of propositions might claim the proposition I find to be ineffable might simply be the necessary proposition. But such a theorist must allow some kind of value to the different ways of specifying that proposition. If this is so, there will be some way of *entertaining* the necessary proposition which cannot be stated since, as before attempts to get at the necessary proposition in *that* way will always fail—they won't even be (simply) true. This, I believe, is interesting enough in its own right to be a substantial conclusion.

I cannot convince myself of this, and do not know how I *could* begin to argue for these kinds of claims. As a consequence I feel that I understand a proposition that is the limiting case of the propositions which, to my mind, are expressible, in the limited cases. Each limited proposition is entertainable by anyone, but only stateable only by some at certain times. In the limit we have the same kind of intelligible proposition that we had in each of the individual cases, generalized in a way to be unstateable by anyone, in any language, at any time. That is to say, we have singular essential propositional ineffabilities in English.

2.5 Clarifications and Objections

With my argument in place, I'd like to clear up some potential misunderstandings of it, and rebut some objections that might seem natural.

First, it might be appealing to dismiss the kind of expressive limitation I've tried to argue for here as familiar and unproblematic. Consider Fadiya, who at time t produces no utterances. Then there is a truth about Fadiya at t that, at t, Fadiya cannot truthfully state. We might state it as I did:

(1) Fadiya produces no utterances at t.

If Fadiya tries to convey this kind of inactivity at t, she will produce an utterance, thereby falsifying her claim. Of course, at earlier and later times she faces no difficulties stating this proposition nor, at any time, do we. But there might be generalizations of Fadiya's problem just as there are of Mika's, which we can arrive at by exploiting the style of reasoning involved in Fitch's paradox. Grant for the moment that there is some proposition, p, which no one will ever state. Then there is a conjunctive proposition which cannot be truly stated.

(2) p and no one will ever truly state p.

If anyone were to state the fact schematically represented by (2), they would thereby state p, falsifying the second conjunct. So (2) may give us the form of a not-truly-stateable proposition.

These kinds of limitations may or may not be interesting in their own right. Importantly, though, they are *not* instances of ineffabilities in my sense—propositional or conceptual. (1) and (2) are cases where one cannot *truly* express a proposition on a technicality. One can't state the proposition truly because, if one were to try, one would thereby ensure the proposition was false.

The examples I've given in the previous section are doubly dissimilar. First, trying to literally express the proposition I'm after does *not* change its truth value. In fact, the proposition in question is a *necessary* truth—*nothing* will change its truth value. Additionally, my case is not one where one cannot *truly* express a proposition. It is one where one cannot express a proposition *simpliciter*. This is all the more vexing because the proposition is a necessary truth. So the cases I've raised, if they behave as I claim, are much more problematic than those given by (1) and (2).

Granting the strength of my conclusion, how might one resist it without giving up (I_1) or (I_2)? A tempting way to do this is to point out the paradoxical or sometimes contradictory statements that I make as I proceed in giving my case. But of course, many ways of doing this encounter a dialectical problem: I am *intentionally* and *knowingly* producing such statements. They are my favored heuristic to help get my readership to think the ineffability which I feel exists. I am happy to withdraw the statements (as I have been conceding). But to point out the difficulties in making my case without some such aid is to beg the question against my position: if there *are* constructible singular essential propositional ineffabilities, then clearly their existence can't be brought out by stating them or deducing them in any straightforward way.

A more promising line of resistance is to try to show that the kind of reasoning that I go through, when generalized, leads to incoherent conclusions, even granting the possible existence of essential propositional ineffabilities. A natural way of doing this is to try to extend the reasoning of the paradox from language to thought. Let me explain why this generalization might seem especially threatening.

I have tried to argue for the existence of ineffabilities constructively: by producing one which is thinkable but unstateable. And the way that I am able to do this is by, in a loose metaphorical sense, treating 'thought' as a metalanguage and all spoken languages as object languages. E-reference is defined for utterances—uses of bits of spoken language. We then reflect on the limitations of the sphere of linguistic use within a separate sphere not up for discussion, the realm of thought. This raises two related worries: First, that the motivations for (I_1) and (I_2) are not accurately reflected in those premises. We should generalize claims about truth to be about *thought* as well as about language. The second, related, worry is that once we do this, we'll see that we were able to entertain my allegedly ineffable proposition only by positing an artificial limitation. Once we remove this limitation, the reasoning will lead to a conclusion that is not merely surprising, but incoherent. Why? Because the conclusion of my argument is that there is a thinkable but unstatable proposition. But if we, say, generalize e-reference and super-reference to thought, by parity of reasoning won't we be able to demonstrate the existence of a thinkable but unthinkable proposition?

Of course I cannot accept this conclusion. But there are several reasons to resist this parity of reasoning argument.

First, there are important potential asymmetries between language and thought which might make a transposition of the paradox harder than it looks. The paradoxes I constructed drew essentially on the idea that linguistic entities exhibit a certain kind of structure. This included the kind of structure required to isolate particular propositional bearers of truth and falsity, and also the kind of structure characteristic of existential form, so that the coherence of talk of satisfaction gets a grip. But there are worries that this kind of structure—a structure characteristic of compositionally generated content in language—is not present in thought. I am skeptical that thought has, or can always be taken to have, such structure. Without it, we cannot properly even *formulate* an analog for the paradox, let alone assess it for coherence.

Since this might be seen as an idiosyncratic defense, let me set it aside for there is a second, structural, problem with transposing my argument to thought. My argument proceeds roughly like this: I begin, relying on two assumptions, to try to draw out the existence of a proposition that can coherently be thought. I then note that this thought resists literal statement in language, and I conclude that there is a thought inexpressible in language. To run the argument again for thought, we would need a corresponding first step: one would have to be able to 'show' a corresponding, coherent thought about the limitations of thought. That first step, however, doubtless can't be executed. The precision and clarity required in the thought to make the argument compelling will be in direct tension with the dialectical purpose of producing such a thought: to produce one which is unthinkable. It's important to note that no such tension exists in my original arguments for essential propositional ineffabilities, so long as we are willing to countenance my coherent, though admittedly surprising, conclusion.

We might helpfully compare my earlier case with Smith here. Smith and I can think, without stating, true answers to my question: "What is an example of something neither Smith nor I will speak of in giving an answer to this question?" What if we generalized the question to be about thought? My new question would be: "What is an example of something neither Smith nor I will think of as an answer to this question?" Does this generalization lead to incoherence—to our being able to think a true answer to the question which is not thinkably-true-by-us as an answer to the question? Not at all. Smith and I simply can't think true answers to the question anymore. (The question, though, will of course still have true answers.) The point is that the generalization of the initial case is not fully analogous to the original one—it raises special new problems in thought. Not only is it *open* to say this might happen for the generalization of my ineffability arguments to thought. Rather, the inherent difficulty in reproducing a first step of the generalized argument seems to show precisely that: the argument simply can't be run in the same way in thought without substantially changing what is taken for granted, and what is argued for.

Finally, *even if* one could rerun the argument in thought without substantially changing the nature of the argument, and its conclusion, it's not clear that what we should conclude from the resulting parity of reasoning argument is that the *conditional claim* I'm arguing for is incorrect. In fact, it seems that were we to rerun the argument successfully, somehow, in a way that led to incoherence, it would be clear that the problem would lie with the analogs of (I_1) and (I_2) for thought. If that's right, then my conditional claim: that (I_1) and (I_2) lead to singular essential ineffabilities might stand undefeated. It's just that we should be using it in a reductio.

As I said, though, transposing the argument to thought leads to significant changes in the nature of the argument. I'm not claiming that we have nothing to learn from trying to transpose my argument to thought at all. What would be suspicious is an attempt to produce a limitation in thought, as I have tried to produce for language, in a *constructive* way. Considerations nonconstructively demonstrating the existence of unthinkable propositions are perhaps worth countenancing. Lewis, on one sense of 'proposition', argues for something like this claim as I will shortly discuss. The resulting view is not obviously incoherent, except on certain specific conceptions of the nature and structure of thought. I don't want to endorse this way of reapplying my style of argumentation—only to note that even those who feel like my reasoning must generalize to thought somehow will not necessarily be led into incoherence by those generalizations.[18]

3 Sources and Implications of Expressive Failure

There are more objections to consider. But I suspect that much resistance to my argument stems from incredulity at the very idea that strong expressive limitations of the kind I'm defending could coherently exist. Consequently I'd like to highlight the features of my ineffabilities that might seem to call out for a special explanation and offer a theory to account for them. The theory, developed in §3.1, is based on the idea that the paradoxes of ineffability arise because of special kinds of compositional tensions. Though it will necessarily be a sketch, and requires more controversial assumptions than I have appealed to so far, the theory importantly gives one example of how we might coherently accept strong forms of expressive limitation and study their linguistic sources. A second reason for supplying this account is that it will shape how I respond to a final objection to my position in §3.2, which will force me to attenuate my conclusion that there are singular essential ineffabilities in English. Even the attenuated conclusion, however, has extremely important implications—most notably for the project of giving a formal theory of truth. I discuss those implications in §3.3, and the equally pressing practical implications of expressive limitation in §3.4.

3.1 How Might Expressive Failures Arise From the Paradoxes?

Historically, expressive limitations have been admitted by a number of philosophers in several distinct domains of study. A quick comparison of three of these domains will be useful as a prelude to further probing the source and severity of the expressive limitations I've been after.

One notable domain in which ineffabilities might appear is in discussions of qualia. If we admit the existence of special 'feels' only available from a first-person perspective, we leave open that information about those feels might be especially hard to communicate in certain circumstances. This might be part

[18] A similar conclusion will be reached for attempted extensions of my argumentation to propositions communicated, instead of propositions thought by an individual. This may be important for the plausibility of my remarks in §3.

of the reason why Jackson's Mary, raised in a purely black and white environment, seems unable to learn what colors look like from the detailed, but colorless, texts in her environment.[19] On some theories of qualia, one can't communicate qualitative information about one's experience unless one's addressee has had similar experiences. On more extreme variants, one can't communicate information about the *very* experiences one is having—this would require distinct individuals to have the *numerically* same experience.[20]

Another domain sometimes alleged to point to expressive limitations in *literal* language is the domain of metaphor. For example, if it is allowed that metaphor is usable to communicate the same kind of information as literal language, and that metaphor is nonetheless not 'translatable' into literal language, then there will be propositional and conceptual ineffabilities relative to the fragment of any language responsible for literal expression. And these ineffabilities could be essential ones if we cannot assimilate metaphorical content into literal language, or at least not all of it. A position of this kind has been advanced recently by Camp (2006) who, if I read her correctly, takes the class of metaphorical contents to constitute a class-based essential ineffability: any individual metaphorical content might be assimilated into literal language, but we cannot overcome any and all need for metaphor in this way.

So far, my ineffabilities are more extreme than those advanced on any of these views. The kinds of ineffabilities arising from qualia are, at worst, *relative* ineffabilities (of the sort witnessed in §2.3) and not *absolute* ones. For every qualitative piece of information, even about a particular 'feel', there is always someone, somewhere, who could in principle express the relevant information.[21] And few discussing metaphor have taken the view that some *particular* metaphorical content is of the same kind as literal content, and yet cannot be integrated into literal language. So qualia and metaphor are not well suited to the production of singular essential ineffabilities.

One of the few philosophers I know of to have explicitly argued for singular essential ineffabilities is David Lewis. His argument, which also implies the existence of unthinkable propositions, draws on elements of the possible worlds conception of a proposition and cardinality considerations concerning the set of possible worlds. It could be paraphrased as follows:

> Grant for the moment that there is some thinker A at time t such that, if a proposition is expressible (thinkable) at all, it is *possible* for A to express (think) it alone at t. Let κ be the cardinality of the set of possible worlds, and suppose there are at least as many propositions as truth conditions—i.e., as $\mathcal{P}(\kappa)$. Now suppose by way of contradiction that every proposition

[19] Jackson (1982).

[20] A related area of potential expressive limitation is in the communication of first-personal thoughts. Frege (1956) defends the view that each person's first-personal thoughts are distinct and unshareable, hence incommunicable—a view with parallels the 'extreme' view about the incommunicability of information about one's private qualia just mentioned.

[21] The same goes for first-person thoughts which, if in fact incommunicable to others as Frege held, should nonetheless be communicable to *oneself* (as in a diary, for example).

> p were expressible (thinkable) for A at time t. If so, there would be at least one possibility where A expresses (thinks) p at t for each proposition p. But since these possibilities are distinct, this would imply the existence of at least 2^κ worlds—a contradiction. So not every proposition is possible for A to express (think) at t. If our original assumption is correct, this implies not every proposition is possible to express (think) at all.[22]

Lewis was not troubled by his ineffabilities since most were obviously too 'irregular' to be worth expressing or thinking to begin with. His argument, however, has a number of questionable assumptions. It is not obviously a safe assumption that there is a single, 'ultimate', collection of possible worlds, or that if such a collection did exist that it would be set-sized. It is also unclear that we should acquiesce in the claim that there is some person and some time such that, to be expressible at all, a proposition should be expressible by that person at that time.

In any event, Lewis' argument brings out one final extreme feature of my ineffabilities: their constructibility. Falling back on the idea that singular essential ineffabilities are 'uninteresting' is what enables Lewis to comfortably adopt his conclusion. By contrast my ineffabilities, if true, concern important limitations of linguistic expression—limitations which are attested to by the very attempts to convey the ineffabilities.

Accordingly, my interpretation of the ineffability paradoxes calls out for explanation of how there could be inexpressible truths with the odd combination of properties I claim for them: propositions, unlike those about qualitative facts of experience, that are inexpressible in literal language by anyone ever (so that they are absolute, not relative ineffabilities); propositions that, like metaphorical ineffabilities, can be conveyed without literal language but, perhaps unlike them, cannot be assimilated into literal language; propositions that are, unlike Lewis's ineffabilities, 'constructible' and fully worthy of expression. I would like to offer one such explanation, even though it will have to be sketchy and go beyond the scope of this paper in drawing on some controversial views about the foundations of linguistic expression and semantic paradox. My hope is that this discussion may at least start to allay worries that there is *no* coherent way of understanding of the paradoxes of ineffability in the way I've argued we should.

I want to begin with a simple idea about the foundations of linguistic expression: that for an utterance to bear content requires some measure of coordination among language users on what content the utterance bears. The idea here is broadly in the spirit of Grice (1957). For "help is on the way" to mean what it does in English requires speakers of English to use, or be disposed to

[22] This is a reworking of an argument considered in §2.3 of Lewis (2001). In fact, Lewis explicitly only rejects the claim that there is some person at a time who could possibly express or think any thought at that time, but his subsequent discussion of that rejection makes clear that his reasons for thinking this are that "[m]ost sets of worlds. . . are not eligible contents of thought." The argument is Lewis' way out of a paradox he attributes to Kaplan.

use, that expression with that meaning in communicative interactions (perhaps by trying to engender certain beliefs). This means that these speakers must coordinate on what propositions those words express in context. If those propositions determine truth-conditions, then coordination on propositional content may equally require coordination on those truth conditions. This kind of coordination needn't be the product of stipulation. It may amount to no more than patterns among the dispositions of speakers and listeners. But without the requisite coordination, however it arises, the expression "help is on the way" could not conventionally be used by English speakers to communicate what it does, a fact which would threaten rob that sentence of its standing meaning.[23]

Compositionally determined content may require even more coordination than this. It may require coordination on the function or purpose of the relevant constituent linguistic expressions in an utterance, and the role they play in the determination of the content expressed. For example, "help" has a certain meaning in English, given by the contribution it makes to the propositional content associated with utterances containing it. "Help" has this meaning *in* the expression "help is on the way" in part because speakers coordinate on the contribution that word makes in utterances of that sentence, and use that contribution to help determine the propositions those utterances express in context. If speakers *never* did this, "help" might not have the meaning that it otherwise does in utterances of the sentence "help is on the way," and, unless it had an idiomatic sense, "help is on the way" might also end up losing its standing meaning as well.

These two preconditions of contentfulness—coordination on truth-conditional content and on compositional function—open up the possibility of two corresponding kinds of semantic defect: coordination failures among language users on the truth-conditional content of an utterance, and coordination failures on the compositional functions of its parts. The idea I'd like to draw on, which is common though controversial, is that the semantic paradoxes bear a form of serious semantic defect: paradoxical utterances fail to conventionally express propositions. Moreover, I'd like to claim that the source of this expression failure owes to the fact that paradoxical utterances exhibit one or both of the kinds of semantic defect owing to coordination failures just discussed.

Now, it's a difficult matter to argue that the paradoxes of ineffability exhibit the defects I've just described, so I won't try to do that here. Rather, I just want to show how the assumption that they do has the potential to explain *both* why my paradoxical utterances don't express propositions, but also why they 'almost' do. The idea is that there is a natural proposition for the paradoxical utterances to express, but they cannot be conventionally associated with that proposition because of their compositional structure. Consider how the paradoxes of ineffability arise: There is an easily recognizable form of linguistic limitation, granting the assumptions (I_1) and (I_2), in our ability to e-refer. We can easily arrive at this conclusion by a chain of reasoning from those intelli-

[23]The coordination, of course, need not always be universal. It suffices, for example, for a suitable range of 'expert' speakers to coordinate on the expression's use, with suitable deference from non-experts.

gible assumptions. The problem is that when we try to state that limitation, the linguistic tools we have for doing so could only succeed in stating that limitation, *with their constituent expressions retaining their standard functions*, if they overcame it.

This fact, I suggest, creates a special kind of 'compositional tension' for utterances of paradoxical sentences like (A). If we assume that its constituent expressions make their normal compositional contributions, the sentence is incapable of expressing the 'limiting' proposition of §2.3 (and indeed a proposition with any truth-conditions at all). This is because the contribution made by some of the constituent semantic words in (A) (like "true") depends, in a degenerate way, on which proposition the sentence is conventionally used to express in the way all liar-like paradoxes do.[24] Which things should be included among those spoken of as "true" or "true of" one another in such paradoxical utterances hangs in a degenerately circular way on which proposition, and what truth-conditions, are compositionally associated with the utterance of (A).

Let me try to characterize this tension in another way. For an utterance of (A) to express a proposition in a normal manner two things would have to occur:

(i) the utterance of (A) would have to be conventionally associated with truth conditions, and

(ii) (A)'s constituent expressions, including those designed to be responsive to distributions of truth-conditions like those posited in (i), would have to compositionally determine in standard fashion truth-conditions for the utterance of (A) which 'line up' with those given in (i).

The particular kind of degenerate self-reference in (A) makes achieving both these tasks at the same time impossible. Whatever truth-conditions are are associated with the utterance of (A) in (i) lead to a compositional determination of truth-conditions, through (ii), that are incompatible with those given in (i).

My first claim, then, is that this compositional tension is enough to ensure that coordination on the truth-conditional content of a relevant utterance of (A) is problematic enough to ensure it does not conventionally express a proposition with truth-conditions. So far this account of semantic paradox at least explains why (A) does not express any propositions, and so *a fortiori* why it doesn't express the proposition I have been claiming is ineffable. But an important question remains: why are we tempted to associate that proposition with (A)—a fact which makes (A) instrumentally useful in ways on which I've been capitalizing on all along?

Let's grant that some steps of the reasoning leading gradually to (A) can be expressed in a conventional manner with literal language. The final statement of the reasoning—(A) itself—by contrast is defective owing to special features of its compositional structure. This nonetheless leaves open that we may arrive

[24]Of course "true" doesn't appear in (A). It would, however, if we unpacked the definitions it contains.

through reasoning at the relevant concluding thought about linguistic limitation in a special way: speakers can successfully associate (A) with the relevant proposition by *ignoring* aspects of the utterance's compositional structure. The idea is that the utterance 'provokes' the thought—as a kind of natural or 'best candidate' proposition to take as being expressed—even though on reflection we may realize that the utterance cannot express any proposition with its constituent expressions retaining their standard meanings and compositional functions. Indeed, this seems to characterize precisely the kind of 'double take' that is a normal reaction to the paradoxes of ineffability. The kind of compositional tension generating paradox only threatens the utterance's *conventional association* with the simple bivalent proposition, but not the truth and intelligibility of the simple bivalent proposition, nor the fact that it is the best candidate proposition for the utterance to express.

So to recapitulate the elements of the proposal: any paradoxical utterance is subject to compositional tensions which preclude it from expressing a proposition with conventional truth-conditions. But the utterance of a sentence like (A) is special because it has an intelligible proposition that is a best candidate proposition to be expressed by it. That proposition still can't stand as the proposition conventionally expressed by utterances of (A) for the reasons already given. But (A) can be used in a kind of derivative way—to springboard interlocutors to a thought by ignoring one side of the compositional tensions in it. We can associate (A) with a proposition only insofar as we ignore the fact that taking this proposition to be conventionally expressed by (A) is at odds with the function of its constituent semantic expressions.

The problem is that not only is the proposition in question a 'best candidate' to be expressed by (A), but (A) is also a 'best candidate' compositionally structured sentence to express that proposition. Any sentence which stands a chance of expressing the proposition will *ipso facto* exhibit the kinds of compositional tensions that (A) does, because it will be forced to talk about itself in degenerate ways. This is the kind of 'linguistic technicality' I spoke of earlier, which has a loose parallel in my discussion of Smith's unanswerable question. If the literal content of a composite expression is determined in line with the meanings of its constituents, *and* our best ways of conveying a particular proposition require defective combinations of such meanings, we can then understand why propositions about distributions of semantic properties can be conveyed without *ever* being stateable.

This account of paradoxical statements like (A) thus provides an explanation of why my ineffabilities are absolute, not relative, constructible, entertainable, and communicable, but not literally statable. I don't pretend to have made an understanding of the paradoxes transparent in this discussion, or my account of them unassailable. But I hope that I have been shown that there are tools available, and worth exploring, that would make much more intelligible the diagnosis of the paradoxes of ineffability that I argued, in §2.4, we are forced to make.[25]

[25]It may seem surprising that I do not defend many striking consequences from singular

3.2 A Final Way Out?

I've sketched the foundational theory above not only because of the need to have some account of the nature and sources of expressive limitation, but because my particular favored account helps to shape my responses to two final objections, which may require my conclusion that there are singular essential ineffabilities in English to be attenuated. A first objection asks: "If we can think the proposition in question, can't we express it with the help of indexical like "the thought I'm now thinking is true"?" A second asks: "If we can think the proposition in question, can't we *stipulate* some new expression to stand for that proposition?" The answer to both questions will depend on whether or not the indexical utterance, or the stipulated expression, could at some level be construed as making existential claims. If they would, then it is easy to see that the new attempted forms of expression fail: the form of the claim would again ensure that it exhibits the problematic self-reference that precludes it from expressing the relevant proposition. The answer to the first question also depends in part on our semantics for definite descriptions, indexicals, and semantic vocabulary like "true". An 'indirect' attempt to avoid the expressive limitation will express the *exact* proposition we want only on some semantics for those terms, and some conceptions of propositions, but not others.[26]

I have no decided opinion on either set of issues, and I want to leave open that they might weigh against allowing singular essential propositional ineffabilities. After all, the account I sketched in the previous section does not speak against the possible success of either strategy. On that account, the expressive limitations arising from paradox are due to special *compositional* tensions—tensions between the purpose and function of an utterance's constituent expressions, and the propositional contents we are drawn to associate with their concatenation. The two approaches above, especially the second, plausibly seem suited to skirt the tensions in question.

To make headway, let's hypothetically grant that one of these two methods allows us to literally express the proposition. What conclusions would we be

essential ineffabilities for the philosophy of language more generally, beyond those for formal semantics and for formal theories of truth as I will shortly do. The reason for this is that on the account of the source of linguistic expressive limitation just given, there are in fact relatively few implications of this sort. I've taken care, for example, to argue for my conclusions independently of any particular conception of propositional structure, or the fineness of grain of linguistic or mental content. Moreover, I see no reason to think that we even need any special alterations to our theories of literal linguistic communication, beyond those which would be required to accommodate something like the view I described in this subsection—the real controversy there again lies in the views presented about the source and nature of semantic paradox. It is in part owing to the fact that the assumptions required to generate ineffabilities are so *weak* that their existence makes less contact with controversial theses in the philosophy of language than one might have expected from what is, on its own, so controversial a claim.

[26]To give one example, suppose we adopt structured Russellian propositions. Consider which Russellian proposition is expressed by a relevant utterance of "the thought I'm now thinking is true". If "the thought I'm thinking now" doesn't contribute a constituent proposition, but a separate structured entity, or if "true" contributes added elements to propositional structure, we may end up expressing a *distinct* Russellian proposition than that we're thinking (though, perhaps, one truth-conditionally equivalent to it).

left with? We would obviously lose the most striking conclusion that English contains singular essential propositional ineffabilities. But we would retain a similar, equally striking claim: that certain propositions exhibit a special, systematic resistance to literal expression with the tools *designed* to express them. That is to say, we can't skirt the expressive limitations by stipulating new conceptual resources for talking about semantics. No relevant concepts of this kind were lacking, and the addition of new semantic terms doesn't overcome our difficulties at all. The only way to circumvent the ineffabilities is by using tools to allude to our thoughts, or by giving piecemeal stipulation of propositional contents one by one (doubtless, with the help of allusion to thought). Moreover, since the range of problematic propositions is likely infinite, and each stipulation removes one ineffability at a time, the stipulative strategy would leave us with a very peculiar form of class-based essential propositional ineffability.

So even if one of the two above strategies succeeds, the compositional tools we have for talking about semantic properties sometimes necessarily fall short of their mark, and no augmentation of our language with similar semantic tools will do the job. This has very important implications for the project which forces us to confront the possibility of expressive limitation—the project of giving a formal theory of truth. Let me now say exactly what those consequences are.[27]

3.3 Consequences for Theories of Truth

We can gauge theories of truth by the trade-offs they allow in three main areas: logic, the behavior of the truth predicate, and expressive power. The arguments I've given place important constraints on how those trade-offs can proceed.

What kind of expressive power do we have reason to want in our theories of truth? First, there is a hankering after the idea that natural language is 'universal' in a sense loosely glossed by Tarski:

> A characteristic feature of colloquial language (in contrast to various scientific languages) is its universality. It would not be in harmony with the spirit of this language if in some other language a word occurred which could not be translated into it; it could be claimed that 'if we can speak meaningfully about anything at all, we can also speak about it in colloquial language.'[28]

What Tarski meant by this is far from obvious, but a few things he couldn't have meant are clear. For example, if universality amounts to the ability to

[27] One more objection might be useful to mention here in passing: can't individuals at 'other possible worlds' express the proposition about our world that we can't express? Wouldn't this make the proposition 'possible to express' in some sense? I'm willing to allow this (though I'm not sure it's clearly true). Discussing this objection in detail would require even more distinctions about kinds of ineffabilities than I produced in §1. What I want to remark here is that even if we grant people in 'other worlds' can express my proposition, this won't affect the important conclusions I'm about to draw from my kinds of expressive limitations for formal theories of truth.

[28] Tarski (1983) p.154.

express every concept and every proposition in some fixed language at some time, there is good reason to think that no human language ever was or ever will be universal in that sense.

A more plausible, but somewhat murkier claim is that we have a presumption that, for any theoretical domain of interest, our colloquial language is able to accommodate sufficient linguistic resources to satisfactorily, if not exhaustively, characterize that domain. As applied to natural language semantics, this would mean that natural language could in principle contain the resources to characterize semantic properties and propositions about natural languages generally to our theoretical satisfaction.

The recourse here to the vague terms "sufficiently broad" and "to our theoretical satisfaction" is necessary. Certainly no one expects English to be expandable to contain a predicate expressing every semantic property. For example, the set of semantic concepts

$$\{\text{refers to } r \mid r \text{ is a real number}\}$$

is too large for any natural language to express piecemeal. What we expect of natural language is something a little hazy, but easy for us to understand: to be able to literally discuss the semantic properties and facts relevant to its operation that are of immediate interest to us 'all at once'. This admittedly imprecise requirement can be split into two parts. First, a requirement on the semantic concepts we can express:

> *Semantic Conceptual Universality*: Natural language can in principle contain vocabulary expressing all suitably interesting semantic properties possessed by uttered sentences of natural language use.

For example, if utterances of any language possess properties akin to truth values among some set $\mathcal{V}$ including, for example, truth and falsity, we would expect there to be a natural language, with a model of it, which could express every concept in $\mathcal{V}$.[29]

A similar analog for semantic propositions can be given:

> *Semantic Propositional Universality*: Natural language can in principle contain vocabulary expressing all suitably interesting propositions concerning the semantics of natural languages.

What my arguments show is that if "true" applies to utterances of all languages in any coherent way, we may have to jettison Semantic Propositional Universality. The ineffable proposition generated on that minimal assumption should qualify as 'suitably interesting' not simply because it happens to be a constructible ineffability (in the way that the claim *that toes are toes* might have been interesting were *it* to have been ineffable). Importantly, the ineffability concerns a structural generality about linguistic use which is part of the

[29] Even this claim might be too strict for some theories. If it turns out the set of truth values is the size of the continuum, for example, we would relax the requirement a bit.

reason ineffabilities arise at all. That the set of all utterances has a unique super referent is a fact about the bounds of actual linguistic expression. The fact that what speakers of any language accomplish is bounded in this way is needed to establish that the essential expressive limitations I've argued for exist. That is it to say, we can't fully understand the *grounds* for the existence of my ineffability—how the ineffability comes about—without, in the process, entertaining the ineffable proposition itself.

So, again on the assumption that "true" applies to utterances of all languages in any coherent way, there are propositions about semantics which are interesting enough for us to want to express, which no natural language of the sort we know can. Even if forced to attenuate my conclusions as per §3.2, my argument will show that certain interesting propositions concerning the semantics of natural language can't be expressed with the compositional semantic tools designed to express them. The only way to express those propositions is by indexical allusion to thought, or piecemeal stipulation of propositional content.

Note that this does *not yet* mean that we should abandon hope for Semantic Conceptual Universality. That would only follow if propositional ineffabilities entailed conceptual ones, or we had a distinct argument from that I have given here that semantic concepts of interest to us cannot be expressed. Accordingly *so far* I have supplied no argument against semantic conceptual universality.[30]

How does this relate to our formal theories of truth? Whether we accept the stronger or the weaker version of my conclusion, we should expect to find special limitations arising in our formal models of natural language use.

> **Conclusion for Formal Theories of Truth**: If a formal theory of truth admits a set of utterances that it is appropriate to speak of with "true", and it abstracts from the resources required for indexical allusion to thought, *there will be some some suitably interesting proposition about semantics that it in principle cannot compositionally express (even through expansions).*

In addition to limiting what can be expected of our theories of truth expressively, this conclusion defuses an objection to theories which exhibit certain strong forms of expressive limitation: one cannot object to a formal theory abstracting from allusion to thought that it makes some semantic propositions inexpressible. Of course such theories may be objectionable on other, related

[30]See Schnieder (2010) for an argument making use of the semantic paradoxes that might be taken to undermine semantic conceptual universality. I am sympathetic with Schnieder's motivations, but I resist his stronger conclusions because his argument rests on an important assumption: that semantic predicates like "true" should, if they report the presence of a property at all, do this in essentially the way that any ordinary predicate like "blue" or "hard" does. This is an idea I think we should reject. I believe that even non-paradoxical uses of "true" which are benignly circular—such as "all utterances (including this one) are true or untrue"—give grounds for thinking that "true" represents in a non-standard way. Though I don't have the space to discuss it here, I believe these utterances tell against the possibility of developing a viable composition semantic theory which associates the truth-predicate with an extension. And I believe that acknowledging this fact may not only afford a means of resisting Schneider's argument, but opens up room for a positive account of how semantic conceptual universality can be achieved. Again, though, these ideas go beyond the scope of this paper.

grounds. One might, departing from my assumptions, object to it for taking there to be complete, coherent norms governing the use of "true". One might object to it for taking truth to be a notion which could sensibly apply to utterances of any language. One might object to it because of the *particular* propositions which it leaves unexpressed, and inexpressible. But if we grant (I_1) and (I_2) and abstract from indexical allusion to thought, then our formal theories *must* admit *some* interesting singular essential propositional ineffabilities.[31] This substantially constrains what we can hope for from a theory of truth, and may substantially influence which theory of truth we should ultimately accept. The question of which theories are favored by this result, however, is one I'll have to leave for another occasion.

3.4 Frege and Communicative Complacency

Having addressed the theoretical implications of expressive limitation, we're left with a final *practical* question: How should we react to the existence of inherent linguistic expressive limitations of the kind I've argued for? We can find implicit answers to this question in the work of an important group of philosophers I haven't yet discussed. These philosophers were arguably forced to take a stand on this issue for the same reasons I face it: by encountering apparent systematic obstacles to linguistic self-description. I'm speaking now of Frege, Russell, and Wittgenstein.

Russell's theory of denoting arguably grew, in part, out of a concern with essential semantic expressive limitations generated by his earlier theories. His reaction to expressive limitations seemed to bear witness to a general distrust of their existence, at least in the relevant semantic domains Russell was investigating at the time.[32] Wittgenstein, on some readings of the *Tractatus*, represents an opposite extreme. He seems to positively revel in, and marvel at potential inherent expressive limitations concerning his notion of logical form, and the doomed attempts to literally express them.[33] But it is Frege who, I think,

[31] The point that such a theory must admit 'some' ineffabilities raises the question of how many ineffabilities of this kind must be admitted. How can my arguments can be generalized? Could we 'generalize' (A), for example, to produce added ineffabilities? Simple ways of doing so would be to conjoin (A) with an uncontroversial truth. More complex ways would try to reproduce my reasoning for distinct kinds of paradoxical utterances. The issues, especially as regards the latter cases, become thorny, and can't be treated here in a satisfactory way. Suffice it to say that I believe that there are many more semantic propositional ineffabilities, but defending their existence is much more difficult to do. (A) is supposed to provide a 'wedge', which would allow for greater receptivity to treating other, more familiar paradoxical statements as connected with propositional ineffabilities.

[32] Russell's distrust of ineffabilities is famously voiced in his preface to the *Tractatus*, where he says of Wittgenstein's methods:

> What causes hesitation is the fact that, after all, Mr. Wittgenstein manages to say a good deal about what cannot be said... His defence would be that what he calls the mystical can be shown, although it cannot be said ... I confess that it leaves me with a certain sense of intellectual discomfort.

[33] I say "on some readings" since the question of whether Wittgenstein was interested in conveying ineffabilities, let alone whether he succeeded in doing so, is quite controversial. Here

should give us our model reaction to the paradoxes of ineffability.

On certain (admittedly controversial) readings of Frege (1951), Frege ran up against precisely the kinds of expressive limitations I've been focusing on when he drew his concept/object distinction. Almost as soon as the distinction was made, it was shown to be subject to what is known as 'the paradox of the concept horse'. The paradox appears to show that virtually everything Frege *states* about concepts are trivial falsehoods, by his own lights, since if we accept the concept/object distinction only objects can be picked out by referring terms in sentences. For example the claim that "the concept horse is a concept" is false since the words "the concept horse" must, by Frege's criteria, pick out an object, not a concept. Frege's ultimate reaction to the paradox is surprisingly curt:

> I admit that there is a quite peculiar obstacle in the way of an understanding with my reader. By a kind of necessity of language, my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept. I fully realize that in such cases I was relying upon a reader who would be ready to meet me halfway—who does not begrudge a piece of salt.

I don't want to endorse Frege's way out of problems with the concept/object distinction, or even that distinction itself. I want to endorse Frege's apparent *attitude* to the expressive limitations he encounters—one associated with a reading of Frege's remarks which is both the simplest and the most controversial.[34] The reading is as follows: Frege concedes that there are important, coherent pieces of information—Thoughts—that he is out to express such that any particular choice of language designed to express those Thoughts is doomed to 'miss their mark'. But rather than view this as a decisive objection to his theory, he regards it as a peculiarity of passing interest. The reason he does so is because the objection does not of itself threaten the coherence of his *theory*, but only his ability to literally express it in language. This is not of itself troubling, since Frege had clearly felt he had other ways of getting his ideas across.

Frege's reaction here is reasonable because of the *kind* of expressive limitation he took himself to encounter on this reading: the limitation is unavoidable, and its practical consequences for communication are quite minimal. These are, of course, precisely the features I've argued actually do belong to the expressive limitations connected with the paradoxes of ineffability. Those paradoxes

a substantial literature has grown up around so-called 'resolute readings' of the *Tractatus*, which try to avoid taking the statements of that work as conveying ineffable insights. See, e.g., Diamond (2001), Kremer (2001) for a pair of representative works. These readings contrast with a more traditional reading of Wittgenstein's work as found, for example, in Anscombe (1971) and Hacker (1986, 2000). Troubles in interpreting Wittgenstein's views on the ineffable often broach issues quite directly related to semantic-self representation and paradox. See Sullivan (2000) for a study in the area.

[34] Incidentally, I also don't want to suggest that this is the best reading of Frege. I'm using it here primarily for illustrative purposes.

seem to present us with an expressive limitation originating from a 'necessity of language'. And, as with Frege, it seems we may overcome the limitations with forms of communication that do not rely exclusively on the compositional mechanisms generating our difficulties. If this is right I suspect that we have little to worry about. The expressive failures are what they appear to be: peculiar defects necessarily arising due to a kind of linguistic technicality. We should not worry about these kinds of expressive limitations in language if we have other means to convey them. But this does not mean that the limitations have *no* importance. On the contrary, witnessing the linguistic expressive limitations is essential to a complete understanding of the structure and limits of literal linguistic communication in a compositionally structured language, and especially how those limits should be reflected in our formal models. As I noted at the outset of the paper, accepting such expressive limitations has the power to radically transform the presuppositions that go into debates surrounding the semantic paradoxes. I believe this even has the power to privilege a very special class of truth-theories, but that is a story for another occasion.

References

G. Anscombe (1971). *An Introduction to Wittgenstein's Tractatus*. St. Augustine's Press.

E. Camp (2006). 'Metaphor and that certain je ne sais quoi'. *Philosophical Studies* **129**(1):1–25.

C. Diamond (2001). 'Throwing Away the Ladder'. In *The Realistic Spirit*, pp. 179—204. MIT Press, Cambridge.

G. Frege (1951). 'On concept and object'. *Mind* **60**(238):168–180.

G. Frege (1956). 'The Thought: A Logical Inquiry'. *Mind* **65**(259):289–311.

H. Gaifman (1992). 'Pointers to truth'. *The Journal of Philosophy* **89**(5):223–261.

M. Glanzberg (2001). 'The Liar in context'. *Philosophical Studies* **103**(3):217–251.

M. Glanzberg (2004). 'A contextual-hierarchical approach to truth and the Liar paradox'. *Journal of Philosophical Logic* **33**(1):27–88.

H. P. Grice (1957). 'Meaning'. *The Philosophical Review* **66**(3):377–388.

P. Hacker (1986). *Insight and Illusion: Themes in the Philosophy of Wittgenstein*. Clarendon Press.

P. Hacker (2000). 'Was He Trying to Whistle It?'. In *The New Wittgenstein*, pp. 353–388. Routledge, New York.

T. Hofweber (2006). 'Inexpressible Properties and Propositions'. *Oxford studies in metaphysics* p. 155.

F. Jackson (1982). 'Epiphenomenal Qualia'. *The Philosophical Quarterly* pp. 127–136.

M. Kremer (2001). 'The purpose of Tractarian nonsense'. *Noûs* **35**(1):39–73.

S. Kripke (1975). 'Outline of a Theory of Truth'. *Journal of Philosophy* **72**(19):690–716.

A. Kukla (2005). *Ineffability and philosophy*. Routledge.

D. Lewis (2001). *On the plurality of worlds*. Wiley-Blackwell.

V. McGee (1990). *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth*. Hackett Publishing.

G. Priest (2006). *In contradiction: a study of the transconsistent*. Oxford University Press, USA.

N. Rescher (2009). *Unknowability: an inquiry into the limits of knowledge*. Lexington Books.

B. Russell (1905). 'On denoting'. *Mind* **14**(56):479–493.

B. Russell (1908). 'Mathematical logic as based on the theory of types'. *American Journal of Mathematics* **30**(3):222–262.

S. Schiffer (2003). *The things we mean*. Oxford University Press, USA.

B. Schnieder (2010). 'Inexpressible Properties and Grellings Antinomy'. *Philosophical Studies* pp. 1–17.

J. R. Searle (1969). *Speech acts: An essay in the philosophy of language*. Cambridge Univ Press.

P. Sullivan (2000). 'The Totality of Facts'. *Proceedings of the Aristotelian Society* **100**(1):175–192.

A. Tarski (1983). 'The Concept of Truth in Formalized Languages'. In *Logic, Semantics, Metamathematics*, pp. 152–268. Hackett, Indianapolis, IN.

M. Tye (1982). 'A note on the synonymy principle of property identity'. *Analysis* **42**:52–55.

L. Wittgenstein (1961). *Tractatus Logico-Philosophicus*. New York: Humanities Press. D.F. Pears and B.F. McGuinness (trans.).