

Semantics for Semantics

James R. Shaw

Draft, please don't cite

What is the point of giving a formal theory of truth? One reason we need such a theory is to act as a compositional semantics for semantic terms like “true”—a theory of how speakers pair content with whole sentences containing that word on the basis of the meanings of their parts. I’m going to argue that in providing such a theory we face an under-appreciated problem. Foundational work in the philosophy of language as well as some basic empirical data require a special kind of reflexivity in uses of terms like “true” that forces their compositional semantic values to be highly non-standard. After describing the reflexivity in question, I’ll argue that it can’t be modeled using extensions consistently with the goals of a compositional semantic theory. If true, this would show that standard model theoretic semantics is incapable of capturing the behavior of an important range of natural language predicates. The result would also rule out a striking range of formal theories of truth as structurally inadequate. After presenting my argument, I explore how our compositional theories should be liberalized to accommodate the peculiarities of semantic terms and extract some lessons from this liberalization for our understanding of the relationship between foundational and compositional semantics, context-sensitivity, the logic of truth, and expressive power.

1 Foundational Semantics and the Compositional Semantics of Semantic Vocabulary

Let me begin by pinpointing an interdependence between foundational semantics and compositional semantics, and a tension it generates.

Philosophers working on foundational questions in semantics are forced to talk about meaning in extremely general terms. For example, a classic project in foundational semantics is to give the conditions under which any pronouncement or gesture counts as meaningful. Another is to say, in some sense, what meanings *are*. The nature of these accounts commits them to an extremely high level of generality. Completed foundational accounts, for example, make pronouncements about every actual candidate assertion from any language.

The generality of these accounts gives rise to an interesting phenomenon. The *accounts themselves* are a form of meaningful linguistic communication that the accounts are supposed to illuminate. So if our foundational explanations of

linguistic meaning are general and successful, they should subsume the speech acts and linguistic tools used in the accounts themselves. In this way, the success of work in foundational semantics turns on the coherence of a broad range of reflexive applications of semantic vocabulary like “means” or “represents”. The accounts must speak partly about themselves.

The reflexivity here is unavoidable even for reductive accounts that try to supply analyses of semantic notions in non-semantic terms. A proposed reductive analysis, no less than any other claim, involves meaningful symbolic manipulation. Accordingly, the analyses themselves must be properly classified using the reduced vocabulary if the notions analyzed in the theory are sufficiently general. In fact, standard *naturalistic* reductive accounts are especially hostage to the worry I’m raising, since if their reductions lack the relevant generality, they won’t allay the naturalistic worries they are designed to address.

The main reason I’m interested in the reflexivity required of semantic terms here is that it generates a special kind of interaction between foundational work in semantics and a separate though related area of research: compositional semantics. Compositional semantics, as I’m conceiving of it, involves the development of formal accounts that illuminate the way speakers are able to attribute semantic properties to whole sentences or utterances on the basis of the semantic properties of their parts.

These two areas of research often overlap in a significant way. A typical example is supplied by belief reports. Belief reports seem to state relations that hold between believers and various ‘objects of belief’. These objects of belief, however, are nothing other than the objects of speech—the assertoric content that foundational semantic accounts are supposed to explain. A compositional semantics for belief reports inevitably imposes constraints on the nature of these objects of belief and assertion. This raises the issue of whether these constraints harmonize with what our foundational accounts tell us the objects of belief are like. As such, the success or failure of a proposal for the compositional semantics of belief ascriptions ends up being tied to the success or failure of distinct accounts in foundational semantics.

Talk using semantic terms like “true”, “means”, or “represents”, however, generates an even more intricate entanglement between foundational and compositional semantics than belief ascriptions. Whereas any reasonable compositional semantics of belief reports must answer to the possibilities given by our foundational accounts, there is a strong *two-way* dependence between compositional and foundational semantics occasioned by semantic terms. Namely:

- (A) Our foundational semantics must use terms to characterize the assertoric content or effects which make assertions significant. And if the foundational work is complete, the terms used in it constrain the properties that can be intelligibly assigned by our compositional theories; moreover,
- (B) every semantic term from the account in (A) must figure in a systematic compositional semantics to ensure the original foundational account is coherent and intelligible—to show the account *itself* has significance.

If either (A) or (B) aren't satisfied, it spells big trouble for our understanding of natural language meaning. For example, if (A) fails so that the properties attributed in our compositional semantics outstrip those explained in foundational semantics, the compositional theory is a mere uninterpreted formalism, and will cease to do explanatory work.¹ If, on the other hand, (B) fails, so that the properties employed in our foundational semantics outstrip those which are given a compositional semantics, we will have no systematic account of the significance of the foundational account itself. Given the account's reliance on unexplained semantic terms, this would threaten to undermine the theory's plausibility, if not its intelligibility.

Before I go on, I need to explain my formulation of (A). I've claimed that foundational semantics will, in part, classify 'assertoric content or effects'. The reason for this is relatively straightforward. Whatever our foundational account eventually looks like, it has to tell us about assertions. It should tell us what makes them meaningful, what endows them with the meanings they have, and what those meanings come to. If it is going to do its job in a suitably informative way, it had better help us partition assertions, or attempted assertions in various ways—for example into those that are meaningful and those that are not. It had also better tell us what distinguishes assertions of sentences with different meanings from each other, which will doubtless be done either by classifying their effects, or by characterizing something else (e.g., their content) which generates the corresponding effects. This is what I mean when I say that our foundational work will involve characterizing assertoric content or effects.

With this clarification out of the way, we can see that the interdependence between (A) and (B) immediately teaches us three important things. First, to give the formalism meeting the demands of (B) we need a formal semantic theory of semantic terms that is more than just a consistency proof for skirting paradox. We must know which features belonging to sentences containing semantic terms make those sentences meaningful, and show how speakers systematically come to associate them with the meanings they have. That is just to say: we have a special need for a compositional semantics for semantic vocabulary. Second, for (A) to be satisfactory, we should expect the compositional theory to inform us about the assertions used in our foundational accounts—presumably by helping us to see how they are “good”, or appropriate, or coherent in the senses we would like the assertions in our foundational accounts to be. Consequently the theory should afford us linguistic tools which partition assertions in the desired way.² Third, the linguistic tools effecting the partitions should be able to do so allowing for systematic *reflexivity* so as to show how the two-way connections drawn with (A) and (B) are possible.

¹ Importantly, it might not *even* allow it to yield a consistency proof, since without the requisite foundational work, the notion of 'consistency' yielded by the formal theory will be unexplained.

²I grant that the way it might do this is 'derivatively'. I don't want to rule out, for example, that a 'transparent' notion of truth geared primarily towards talking of non-linguistic facts could nonetheless indirectly give us information about the linguistic entities which speak of those non-linguistic facts.

It is natural, if one works on the paradoxes, to be suspicious of the kinds of general requirements on formal theories of truth (and other semantic terms) that I am tracing from our foundational goals in the philosophy of language. After all, many have claimed that the paradoxes teach us that much more familiar and cherished desiderata must go by the board. So why think that standard work in the philosophy of language should shape our approaches towards truth. Why not think, instead, that our work on truth will give grounds to abandon the projects I've been alluding to?

I'm open to this possibility, and one thing I want to draw attention to is precisely that the effects on most foundational work in the philosophy of language is often ignored by those exploring various forms of expressive and logical restrictions in their formal theories of truth. The foundational programs that I've been alluding to won't, in any recognizable way, survive weakenings of the assumptions I've been making—particularly the assumption of generality. If the semantic paradoxes teach us that the reflexivity required for (A) and (B) is not coherent, then we should, I suspect, abandon hope of resolving the problem of intentionality, especially in any naturalistically satisfying way. We should probably even abandon hope of minimal reductive accounts of linguistic intentionality in terms of mental intentionality (or vice versa). Intentional, representational, and indeed linguistic phenomena on the whole must, in a way unlike any other phenomenon we know of, remain essentially inexplicable. Moreover whatever formalisms we give in supplying our theories of truth will be essentially uninterpretable, or open ended, in the sense of lacking a justifiable theory drawing explicit ties between the formalism and the mental and linguistic phenomena that the formalism is clearly supposed to somehow illuminate. This threatens the whole point of giving any formalism in the first place.

Again, we have grounds to think the paradoxes teach us very surprising things, so I don't want to presume that the failure of most foundational work in the philosophy of language, and the failure of the utility of certain standard modeling techniques, aren't among them. But even if this is a conclusion we should *eventually* draw, for now I want to proceed to investigate the interaction between (A) and (B) in abstraction from the challenges posed by the semantic paradoxes in order to see what we can learn independently of them. I'll argue that this leads to some interesting and surprising results. In particular, if we provisionally ignore paradox, the task of structuring a compositional semantics to achieve (A) and (B) on its own requires non-trivial emendations to those theories which are largely unappreciated. In addition to driving us into new ways of thinking about linguistic meaning and representation we may also learn something about other aspects of truth, including semantic paradox, indirectly.

2 Truth-Extensionalism & Compositional Circularity

Before I continue I need to make two preparatory remarks.

First, I want to flag how I will be using the word “property”. In the previous section I noted that our foundational semantic accounts must inform us about assertions and their effects generally: they will give us means of *partitioning* assertive acts through assertion. And to respect the interaction between (A) and (B) the assertions effecting the partitions should at least sometimes include themselves among the partitions. I need a term to talk about the sets of speech acts, or utterances, partitioned in this way and speaking as if they ‘have a property’ seems like a helpful way to do this. But I want to stress this is a special piece of terminology. We needn’t think of these properties in any substantive way beyond that they involve classifications. My use of “property” is thus meant to be compatible with what some deflationists say when they deny that semantic terms track properties. Often this is said to claim, for example, that truth has no ‘nature’, or there is no ‘one thing’ in virtue of which truths are true. This is compatible with “true” playing my meager classificatory role, and for there to be something like a set of truths.

Next, a methodological remark. I’ll be proceeding on an assumption that the English word “true” is among those that records a property of assertions of high interest to foundational semantics. This might not be right. Maybe our use of “true” is simply inconsistent, or expresses a concept so ‘deflated’ that it is unsuited to the kind of important foundational work I want to set it to. That’s fine, and needn’t affect the importance of what I’ve said in §1. If “true” doesn’t do the work we want, we will need to appeal to other notions in our foundational theories—perhaps completely new notions we introduce for the purpose. Then those will be the terms of greatest interest to philosophers of language and logicians. What I’ve claimed in §1 is that we need some terms or other to express properties of assertions (in my sense) if we’re to hope to execute standard programs in the philosophy of language. So for reasons of intuitiveness and simplicity I will proceed as if “true” is one of them, though my main arguments will apply to whatever semantic notions end up being foundationally significant.

With those preparatory remarks out of the way, what we’ve seen so far is that, even if there were no semantic paradoxes, we would need a compositional semantics for words like “true” in a more pressing way than we do for “tall”, “might”, “wrong” or even “believes”. A formal compositional theory for semantic words isn’t just an explanation of how we consistently coordinate on the assertoric content generated by some expression among others. It isn’t even just an account of this kind for an expression whose compositional behavior resists easy systematic explanation, or one which occurs frequently in philosophical discussion. Rather, it constitutes one of two integrated theories required to secure the possibility of a general study of meaning and communication at all. So it is important to ask: what will the shape of this formal account look like?

If the semantic terms in our foundational accounts end up recording properties of assertions in the way I’ve been suggesting, a natural way to model the compositional effects of words like “true” suggests itself. In context, we should assign it the familiar kind of semantic value borne by any common adjective: an extension consisting of the entities which bear the property it is used to record. I’ll call this view “Truth-Extensionalism”.

Truth-Extensionalism: Our compositional semantics for “true” should assign it an extension as its semantic value (perhaps relative to a context/index pair, or along with an anti-extension, etc.).

The next thing I’d like to do is argue that Truth-Extensionalism is false. The reason for this is a conflict between *some* kinds of reflexivity required of uses of the truth-predicate on the one hand, and the most basic goals of a compositional theory of meaning on the other. Before we see how the interdependence from §1 contributes to this conflict, it might be helpful to look at a simpler empirical phenomenon which generates the relevant kind of problem as well.

Suppose Sid, a compulsive liar, has promised Marie on Sunday that for all of Monday he will take care to speak only the truth. On Monday one of Sid’s friends asks about his uncharacteristic honesty, and he explains:

- (1) I made a promise yesterday which I’ll have kept just in case everything I say today is true.

Ordinary speakers report that (1) is unexceptionable: a virtuous assertion along any dimension you could think of, including being true. What’s interesting about (1) is the way in which it speaks about its own status. In a way familiar to truth theorists Sid’s utterance is ‘about itself’, albeit conditionally. Suppose Sid continues in a natural way, proud of his care and honesty.

- (2) Everything I say today will be true.

If Sid otherwise only tells unambiguous truths, speakers report that (2) is again, unexceptionable and true. This is so even though, to all appearances, (2) speaks of, and classifies, itself directly and not hypothetically. Importantly, speakers recognize that (2) is distinct in meaning, and claims more, than sentences like (3) and (4), where *t* is a hypothetical time at which Sid might utter (4).

- (3) Everything I say today, except maybe this, will be true.

- (4) Everything I say today except maybe what I say at time *t* will be true.

This seems obvious: for (2) to say what Sid wants it to, it *must* be about itself. That’s the only way for (2) to state that the exact conditions of his promise will be fulfilled. It’s accordingly the only way for (2) to have the kind of generality Sid intends it to.

Speakers’ reactions to cases like (2) are interesting in that they exhibit a pattern which I’ll call *defaulting*. This occurs for certain utterances which self-ascribe “true” in a such way that (loosely speaking) their truth would contribute to their being true, and their falsehood would contribute to their being false. That is, supposing we think of “true” as bearing an extension, then if the utterance of (2) were in the extension of “true” (along with the other utterances of the day), we could recursively settle on that basis that the utterance should come out true. If, on the other hand the utterance of (2) were not in the extension of “true”, we would settle the opposite. Thus, in some sense, utterances like (2)

can consistently and coherently have either truth-value, but they conventionally ‘default’ to the value *true*.

Before I say more: should we take speakers’ intuitions about cases like (2) seriously? Are they providing us with genuine and reliable information about the semantic status (2) has owing to its literal linguistic meaning? Absolutely. Not only is there a broad and quite stable speaker concurrence on (2)’s assessment, but we have added *theoretical* grounds that bolster those assessments. If we refuse to count Sid’s (2) as true, we generate a special, and perplexing, form of expressive limitation. Sid can easily state truly that the conditions of his promise will be fulfilled on the day of his promise to Marie. He can also state this truly on Tuesday. If we deny that Sid’s (2) could be straightforwardly true, we deny Sid the most obvious means of stating that the conditions of his promise will be fulfilled—but only on the day they *will* be fulfilled. Whence this very selective expressive limitation? The explanation can’t be that what Sid is *trying* to say isn’t true. We can grant that he truly said it the day before, and will truly say it the day after. We just have to deny his ability to say that very thing Monday (and with the words that seem best suited to say it). Moreover, discounting (2) raises a series of additional awkward questions: Can Sid truly state, on Monday, “My promise will be fulfilled”? If not, why not? If so, shouldn’t this, along with (1), entail Sid’s (2)?³

So we have about as strong a case for (2)’s truth as we ever have: strong and stable speaker concurrence, and even important theoretical backing for the relevant speaker assessments. These defaulting cases may admittedly seem somewhat marginal. But the phenomenon they exhibit is of very high theoretical interest. For example, if any utterances do default, this implies that the meaning of the truth predicate isn’t exhausted by the Tarski biconditionals. That is, if a complete understanding of the meaning of the English word “true” involves having the understanding that ordinary speakers do—that Sid’s (2) is true—that understanding requires more than understanding the Tarski biconditionals which, on their own, cannot give us guidance as to what truth value Sid’s utterance should bear.

That makes defaulting cases of more than passing interest. More importantly for my purposes, though, defaulting cases also tell against Truth-Extensionalism. Allowing “true” to bear an extension as its semantic value in this case inevitably yields an unacceptable story of the source of our knowledge of the truth of Sid’s utterance. The problem here arises from the most basic goal of a compositional semantics. A compositional theory of meaning is supposed to provide us with an account of linguistic *productivity*—the phenomenon whereby speakers associate meanings (at least in the broad sense of assertoric effects and statuses) with an indefinite range of utterances they have never heard before. The way it does this is by associating compositional contributions or semantic values with syntactic constituents of sentences and rules for composing those values that

³Should we try to escape these problems by denying that Sid can *literally* express his claim with (2), and shift the burden to our pragmatics by allowing that he can imply it? No. Sid could have equally well promised that what he said *or* implied would be true. Shifting the burden to our pragmatics is only temporarily moving the bump in the rug.

mirror syntactic operations of composition in the sentence. The idea is that speakers can learn the finite stock of semantic values of the syntactically simpler expressions and the rules for their composition, and use them to determine the assertoric contents and effects of an indefinite range of syntactically composite utterances.

If speakers are right in saying (2) is true, they're clearly able to settle this in part owing to their linguistic competence—in part owing to their understanding of the meaning of “true” and how it contributes to (2)'s truth along with other worldly facts. Our compositional semantics owes us an account of the contribution that comes from their linguistic competence.

Suppose that we assign “true” to an extension in (2) as the contribution that “true” makes to settling (2)'s truth-value (relative to the context of Sid's speaking, and the world he is in) by normal model theoretic means. Then how do *speakers* use their linguistic competence, so modeled, to arrive at the understanding that Sid's utterance of (2) is true—that this particular concatenation of words has that particular assertoric status? To determine whether an utterance of “All *F*s are *G*s” is true when *F* and *G* have extensions as semantic values, one settles whether the extension of “*F*” is a subset of the extension of “*G*”. This means one can't settle the value of “All *F*s are *G*s” until one has antecedently ascertained enough about the extensions of “*F*” and “*G*” to give a verdict on whether the subset relation holds. Standard compositional rules tell us that an utterance of (2) is among the utterances which one must know is, or is not, in the extension of “true” for the appropriate subset relation to hold. Consequently, to *determine* what (2) is saying, and that it is true, you must *antecedently know* it is true.

This seems implausible as an explanation of what is actually happening. But more than this, it undermines the whole purpose of the compositional semantics: to explain how you arrive at the assertoric content, effects, and status of a composite expression like (2). The account only explains how you compositionally arrive information about its status on the assumption you already have that very information *non*-compositionally. This is just to say, the account doesn't contribute to an explanation of linguistic productivity here.⁴

Assigning “true” an extension as part of its semantic value makes (2)'s truth depend on its truth in a way that makes the account of its semantic status circular—but circular in a special *way*: it obscures how we could acquire knowledge of membership of certain utterances in the extension of “true” when we do. This is why, if “true” means what it normally does in (2), we must reject Truth-Extensionalism.

To understand what I mean when I say that (2) is circular in a special

⁴It might be useful at this point to flag a potential misunderstanding of what I'm doing. I'm claiming that Truth-Extensionalism is false because it violates constraints on a compositional theory of meaning. But I'm *not* claiming that Truth-Extensionalism violates *compositionality*—the claim that meanings of wholes are determined on the basis of the meanings of their parts. In fact, most of the semantics I'll be criticizing are *clearly* compositional (on many ways of specifying what that comes to). I'm claiming that any Truth-Extensionalist account fails to explain *productivity*. Compositionality is one constraint on a theory typically aimed to ensure it is productive, but it's not the only one.

way, it might be helpful to contrast cases involving kinds of circularity which don't raise any problems either for the exhaustivity of the Tarski biconditionals or for Truth-Extensionalism. The empirical circularities most familiar from discussions of the liar paradox are of this relatively innocuous kind. Consider, for example, typical 'interdependent' utterances of the kind familiar to readers of Kripke (1975) like (J) and (N).

(J) Most of what Nixon says about watergate is true.

(N) Everything Jones says about Watergate is false.

One thing Kripke noted about such examples is that although utterances of (J) and (N) by Jones and Nixon respectively can be paradoxical, under certain favorable circumstances they are not. For example, if Nixon has only uttered several clear falsehoods besides (N), and Jones has only uttered (J), then Jones' utterance is false and Nixon's is true.

One might think that such utterances are 'self-dependent' in ways similar to defaulted utterances. For example, (N) ascribes a semantic property to (J) which in turn applies such a property to (N). There is a sense in which (N) is, in part, 'about itself'. But actually none of this, on its own, casts any doubt on Truth-Extensionalism. After all, every normal case where (J) and (N) are truth-evaluable is one where we can independently settle by transparent compositional means *enough* semantic properties of utterances to first settle the semantic status of one utterance, and then use this in settling the semantic status of the other. Letting "true" bear an extension in familiar forms of composition in this case isn't problematic precisely because speakers are always in a position to work in line with a Truth-Extensional theory to fix truth values of (J) and (N) sequentially. There is no point at which an apparatus assigning "true" an extension would require speakers to have information about (N)'s semantic status to *fix* its very semantic status. Likewise for (J).

Thus even though successful utterances of (J) and (N) may 'carry information' about themselves indirectly through their interdependence, that interdependence does not generate the special kinds of problems for Truth-Extensionalism that defaulting does. Only utterances which, in some sense, genuinely 'rely' on their own truth values can generate the problem I'm pointing to. Truth-tellers, like (T), which are another familiar kind of example in the literature, have the right structure to do this.

(T) The first line labelled "(T)" in this paper is true.

But it is easy to reconcile truth-tellers with Truth-Extensionalism because they are typically also claimed to be defective in some way. If, for example, we don't see fit to call an utterance of (T) true or false, maybe the explanation is just that standard compositional mechanisms fail to engage in it, or at least fail to engage in the right way. If so, explaining (T)'s defects needn't require special semantic values and it becomes possible, on this story, that "true" bear a humdrum extension as semantic value (perhaps, e.g., allowing for gaps).

When we have the behavior given by (2), on the other hand, we need a special explanation of why direct self-reliance sometimes *is* acceptable (and now also why it is *only* sometimes acceptable). We need to know why familiar, non-defective forms of assertoric content get associated with (2) as a function of the meanings of its constituent expressions. This explanation thus must appeal in part to the semantic value of “true” and, as we’ve seen, setting this as an extension won’t yield any satisfying explanation of the productive, compositional assessment of (2).

Since defaulting is an outlying phenomenon it might be tempting to discount it to avoid problematizing our most simple semantics for words like “true”. In the context of building a compositional theory, this seems methodologically backwards. First, our linguistic accounts here are descriptive: they owe us an account of what speakers are actually doing. If, as I’ve argued, speakers’ assessments here are stable, and reflect the semantic status of defaulters, our theories can’t overlook this for the sake of simplicity: the truth is just complex. Moreover, even if we were only interested in semantic revisionism we should be interested in capturing this phenomenon: since there is so much difficulty in giving a complete semantics for “true”, we have every reason to look for stable clues and constraints whose investigation might provide us with insights and unexplored avenues.

But even if one pleads to overlook defaulting, there is another phenomenon which sometimes generates the same problem and which simply cannot be set aside. This is our need to produce *semantic generalities*: statements about broad distributions of assertoric statuses or effects which are so sweeping that they end up capturing themselves in the process. Here are a few examples which, again, seem to generate the special kind of circularity which interests me.

- (5) If a statement is true, then it’s true to predicate “true” of that statement.
- (6) A statement of the form “An F is G just in case H ” is true just in case every object satisfying “ F ” is among those satisfying “ G ” if and only if they satisfy “ H ” as well.
- (7) For all speakers S , propositions p , and utterances U , S meant that p in uttering U if and only if . . .

Many philosophers of language and logic have endorsed something like (5). Let’s suppose, just for the sake of argument, that they are right—(5) is true. Furthermore, let’s call the claim I just made, that (5) is true, “(5T)”. Then speakers can know the truth of (5), and the truth of my claim (5T), on the basis of their linguistic competence along with knowledge of other facts. But how? Our compositional semantics should again contribute to an answer.

But there’s a potential problem. If (5) is true, then (5) and (5T) together form an instance of a general semantic phenomenon about the use of “true” in English which (5) is designed to truthfully record. I’ve said speakers who settle the truth of (5), and the truth of (5T), do so by recognizing the meanings of their constituent words and how those meanings contribute, along with actual

facts, to their truth. But two of the facts on which the truth of (5) hangs, if the meaning of “true” is an extension, are (i) the semantic status of (5) and (ii) the semantic status of (5T). It’s a contingent fact that these semantic statuses align—owing to contingent facts about what various words in English mean, for example. If they failed to align (5), meaning what it *now* does, would no longer be true.

Speakers’ knowledge of the two semantic statuses, and that they align, is supposed to be the *product* of an application of speakers’ linguistic competence, including their understanding of “true”. But if Truth-Extensionalism is correct, that knowledge is also an essential *requirement* of the application of that competence to settle the status of (5). That is, speakers must have available information about the output of compositional processes—the processes that yield the statuses of (5) and (5T) on which they align—for antecedent use the very compositional processes which are supposed to help settle those outputs. If speakers don’t have that ‘pre-compositional’ knowledge of the status of (5) and (5T), a model making use of extensions doesn’t give us a story about how speakers know (5) is true. This is to say, Truth-Extensionalism again doesn’t give us an adequate explanation of linguistic productivity.

The result here is what might be called a ‘compositional circularity’. This occurs when a theory forces us to say that speakers get compositional knowledge of the semantic status of an utterance *U*, when they get it, only by having antecedent non-compositional knowledge of that very status. Compositional circularities are obviously problematic: they tell us that speakers never have need of compositional processes that they in fact clearly do need. You need, for example, to know the meanings of the parts of (5) and facts about English syntax to settle that it is true. So theories with compositional circularities must be rejected.

I’m claiming that compositional circularities arise on any Truth-Extensionalist account of the status of (2) and (5)—though on (5) of course it’s occurring on a ‘grander’ scale. And similar kinds of broader circularities arise for utterances like (6) and (7), which are again recording suitably general semantic phenomena of which their own semantic statuses constitute instances.

The most recent utterances that I’ve considered—(5) through (7)—have a special importance. They represent a very small sampling of the kinds of semantic generalities that are produced in standard discussions of philosophers and logicians with interests in truth, inference, and intentionality. Given what I’ve said in §1, this is no accident: we *must* be able to meaningfully say some such things if a satisfactory theory of meaning is possible. If this is right, we have to reject Truth-Extensionalism not just because it gets the case of defaulting wrong (which already may be grounds enough), but because it stands in the way of our understanding the expressions required of any systematic and fully general theory of meaning at all. At least, this is what I’m contending. To help clarify my case, let me give a few comments about the argument, and how it works.

First, it’s important to note that not just any circularity in a semantic generality could constitute what I’m calling a *compositional* circularity given Truth-

Extensionalism. We already saw this point as regards ‘semantic particularities’, as opposed to semantic generalities. Kripke’s ‘particular’ semantic interdependencies, for example, do not generate compositional circularities when paired with Truth-Extensionalism. Defaulters, by contrast, do. The same division holds in the general case. Consider the semantic generality “every utterance is true or not true”. This utterance, if true, is in part about itself. But it doesn’t lead to a compositional circularity. Speakers might not need to know the truth of any utterance to settle this utterance is true, even if all there is to the semantic value of “true” (at a context) is an extension. For example, speakers might be able to settle just from the meaning of negation that whatever extension “true” has this utterance will come out true. So Truth-Extensionalism needn’t lead us astray in this case. (5) is different: its truth depends on particular outputs of compositional processes—and the output of the process involved in settling the values of (5) and (5T) included. In this case, which utterances count as “true” really matters to whether (5) is true—and (5) is among the utterances that so-matter. That is what makes this case problematic for Truth-Extensionalism.

Second, the problems (5)–(7) raise for Truth-Extensionalism don’t rely on the idea that speakers always or even most of the time, use standard compositional mechanisms to ascertain their assertoric effects. We have all sorts of ways of learning what a composite expression means. One is simply to be told. For example, suppose someone tells me “the next thing Paul will say means “when the upper-crust does shady deeds, they do them all over town””. Shortly thereafter, Paul produces a string of sounds in a language barely recognizable to me. In this case I may know the meaning of Paul’s utterance, and perhaps that it is true, but I clearly won’t have figured any of this out compositionally. I probably won’t even know what the ‘parts’ of the expression are. Another way to determine what something means, or whether it is true, is to infer this from other information—usually semantic—that you have about the utterance.⁵ Speakers may use these sorts of methods, especially the latter inferences to and from semantic information, to discover the status of (5)–(7). These alternative pragmatic means of getting at the semantic status of an expression are, however, parasitic on the existence of, and hence distinct from, the standard compositional mechanisms generating that status. You can only reliably *infer* what semantic status (5) has if something *determines* such a status. And if the apparently ‘indirect’ means of establishing an expression’s assertoric content turn out to be necessary to ascertain it, then it is those methods which should be built into our compositional theory. What we’ve seen in (5)–(7) is that extensions cannot account for the default, semantic process that pairs utterances of those expressions with assertoric content, and this is why Truth-Extensionalism is false.

Third, it’s important to keep in mind is that even if you doubt that (5)–(7) are truthful, or even meaningful, this won’t skirt the problem that I’m raising, which has little to do with these particular examples or even truth itself. I said in §1 that an interaction between (A) and (B) requires us to have linguistic

⁵See Lasersohn (2006) for some helpful examples and a discussion.

tools to assertorically partition assertions according to their statuses and effects in reflexive ways. My point is that whatever tools we avail ourselves of, the generality of our foundational programs requires us to make assertions just like those in (5)–(7) with those tools. Those assertions will generate the special kind of reflexivity which is in tension with treating the tools as having extensions in a compositional theory. So Truth-Extensionalism—and really any interesting form of semantic Extensionalism—should ultimately be rejected.

A final worry one might have about my argument has to do with its consequences. If Truth-Extensionalism were to fail, it is tempting to think we’ve not just shown that “true” can’t have an extension as its semantic value in a compositional theory of meaning. Maybe we’ve shown, by reductio, that “true” can’t have an extension *full stop*—that there is no set of truths (or utterances appropriately partitioned with “true”). Won’t this force us to give up a key starting assumption of my arguments: that “true” and other foundationally significant semantic terms are used to talk about properties of assertions? This would be disastrous, since it would undermine the starting point which I’ve claimed gives us very special motivations to look for a compositional theory of semantic terms.

But fortunately this is an overreaction. To see why, we need to realize that the problems raised by defaulting cases and certain semantic generalities for Truth-Extensionalism have a common and readily identifiable theoretical source. Suppose, as seems natural, we take truth to be a property an assertion has depending both on how the actual world is, and also depending on linguistic convention. In particular, the truth of an assertion is dependent on the (explicit or implicit) conventions establishing the meanings of the words in that assertion and how they interact, perhaps with circumstance, to give the assertion the semantic status it has. Then it is actually not too difficult to see why Truth-Extensionalism would be false. Although truth is a property, and there is accordingly something like a set of true utterances produced in assertion, *exactly which* set of utterances is true depends on the linguistic conventions we use to establish the meanings of our vocabulary. The vocabulary we use to talk about assertion, and its virtues, is no exception. We can grant that no matter what conventions we actually use to establish the semantic value of the word “true”, there will be a set of truths. But which utterances are in the set will depend in part on which linguistic conventions we choose for the word “true”.

What this means, though, is that to try to establish how we use the word “true”, drawing on the inevitable existence of some actual set of truths T , by saying: ““true” is appropriately used of all and only the utterances which are in T ” or “let “true” speak of the property *truth*” may not produce an acceptable definition—at least not one that will illuminate the semantic talk of speakers in the way that a compositional semantics should. The property of truth is not ‘out there’ in any determinate enough way (i.e. it is not yet determinate what the of actual set of truths is) until it is determinate what “true” means—what its compositional effects are. The instantiations of the property depend on the word used to record the property in a non-trivial way. So the way we use the word “true” cannot *simply* be modeled by an extension. That would

be to presume, falsely, that it is in principle possible, with enough worldly information, to identify all instantiations of the property that speakers can *without* the information given by the stipulation of how to record it. But with “true”, the stipulation for its use *constitutes* part of the worldly information which settles when the property is instantiated. If this weren’t the case, it’s not obvious there would be anything peculiar about “true” or any other semantic term.

What this teaches us is that although Truth-Extensionalism is false, the failure of that view has nothing to do with whether or not we use the truth predicate to track a property—something like a set of assertions. The problem with taking this set as the semantic value of “true” is that it distorts the process whereby the truths are generated: it sets the output of the process of compositionally assigning assertoric statuses to sentences containing “true” as an input that contributes to speaker assessments using those compositional processes. The problem is that there is occasionally an in-principle bar to enough of the extension being ‘available’ for speakers to settle semantic statuses in conjunction with compositional processes.

So we don’t have reason yet to think that there isn’t just a simple set of truths—a simple set of well-made assertions along a certain important dimension. We only have grounds to think that this set can’t always be drawn on for use in the kinds of linguistic conventions which will, perhaps inevitably, generate such a set. If my arguments are sound, understanding how semantic predicates work surprisingly requires us to keep a strict *separation* between projects in foundational and compositional semantics, despite the fact these projects are intertwined. In particular, *we must not confuse the question of what the most important terms in our foundational semantics track with the question of what compositional mechanisms we have available to do the tracking.*

To many, especially those who have an antecedent sense for the peculiarities of semantic terms, it may come as small surprise that Truth-Extensionalism fails. Even if this is so, the failure in question has an under-appreciated consequence. To see this, return to the idea with which I began this section: we need a compositional semantics for semantic terms—a formal theory which pairs utterances containing semantic terms with assertoric content. Though, as noted earlier, I want to provisionally abstract away from problems connected with paradox, it is impossible to ignore the many formal semantic theories that have been developed in projects on the paradoxes, each with its own very unique structure. If we are looking for a compositional semantics for semantic terms, one might think that a first place to at least look to for inspiration would be in the various sophisticated semantic theories developed to cope with the liar paradox. These formal theories, whatever their motivations, at least have the right *form*. They pair values, sets of values, or worlds with whole sentences or utterances as a function of their compositional make-up. Even if a theorist supplying the formalism has no aim of interpreting those values in the way a compositional semantics requires, we might avail ourselves of it nonetheless.

So the question arises: which formal theories can we rule out as unsuited to our needs? Which formal theories of truth do not meet the bare minimum

requirement imposed by the denial of Truth-Extensionalism? Some notable theories which use extensions in their formal semantic theories, and therefore cannot be of help, include broadly Tarskian approaches, and the formal semantic theories found in Kripke (1975), Glanzberg (2004), Priest (2006), Field (2008), and Beall (2009). And of course these are just notable representatives. Really, any theory that makes use of standard model theory has to go by the board. I know of only a few theories—for example, the Pointer Semantics of Gaifman (1992, 2000) and the Revision Theory of Gupta & Belnap (1993)—that even *could* have the right structure to meet the requirement.

I bring up this fact because I want to stress that although the idea that Truth-Extensionalism should be abandoned might seem very natural to some, it is not an option that gets very much attention in the formal setting. Now, it's important not to overstate this point either. Many of the authors I mentioned have no interest in providing a compositional semantics of the kind I've been claiming we need. Field and Beall, for example, are very clear that their formal semantic theories should not be aiming to provide a semantics for (a fragment of) any natural language. Instead their semantics are meant to characterize inferential relations or provide evidence of non-triviality.⁶ Accordingly, it isn't fair to say that the failure of Truth-Extensionalism shows that any of these authors has failed in a project they were undertaking, or even a project that they regard as viable. A second reason not to overstate my conclusion is that although a given theory of truth, like a broadly Kripkean semantics, might have to be rejected as a viable compositional theory along with Truth-Extensionalism, there may yet be *reworkings* of that very formal theory which preserve its allotment of semantic properties while carefully working around the use of extensions. The problem with the semantics, after all, lies not in what properties it assigns to sentences or utterances, but in *how* it assigns them.⁷

These caveats don't undermine the main point, though, that little attention has been paid to the need for a formal theory of truth *as* a compositional theory, and the special constraints imposed on that theory's structure by the broad reflexivity required of semantic terms. The result is that we don't have any real sense for what an adequate compositional theory for semantic terms really looks like. So I'd now like to turn now to investigate how we should restructure our theories if Truth-Extensionalism does indeed fail.

⁶Field (2008) p. 356 for example, and Beall (2009) pp.56–7. As I mentioned in note 1 though, it's a tricky matter to say how a formal theory can 'just' supply a consistency or non-triviality result, or a characterization of valid inference, without substantive, explicit ties between that formal theory the intentional properties of the language it is clearly meant to illuminate.

⁷It's not unreasonable, for example, with a Kripkean theory to take the meaning of "true" not to be given by a given fixed-point model, but by the process which leads to that model. This is fine, and may meet the constraints I'm about to propose for a compositional theory shortly. What matters is that the proposal, and its structure, are brought out explicitly, since there are more constraints on a theory of meaning than just productivity. Any formalism in the business of giving us a compositional theory needs to be set out *as* a compositional theory so it can be properly assessed. This is partly what I'm trying to stress that we need to do.

3 Procedural Semantic Values

Let’s remind ourselves how we got here. Foundational semantics requires us to record properties of assertion. But standard predicate denotations—extensions—can’t do the job for the empirical and theoretical reasons I’ve overviewed. A hypothesized set of extensions for semantic terms may constitute a consistent, pleasing, and even ultimately *correct* allotment of properties among utterances. But extensions simply cannot model the conventional, rule-governed method speakers have to associate utterances with the properties allotted when they do so. Consequently what we need to do is *liberalize our conception of predicate denotations* if we’re going to understand how semantic terms work.

The crucial mistake in using extensions for the denotations of semantic terms was that they sometimes required speakers using compositional methods to determine the status of a reflexive utterance to antecedently have information about that utterance’s very status. So when we retool our semantic values for words like “true”, they should be tailored to ensure that the status of reflexive utterances of semantic terms can be compositionally determined without antecedently drawing on information about the *actual* status of those utterances.

I don’t want to give the gritty details of a particular formalism here. This would be tantamount to endorsing a particular formal theory of truth and the associated controversial commitments about paradox that I want to avoid. Instead, I just want to motivate two features that I think should play a crucial role in the way our new semantic values operate. These two features will do the work we need of explaining how the tension between foundational semantics and compositional semantics can be resolved, without making any direct pronouncements on controversial uses of the truth predicate.

The primary thing we need of our new semantic values is that they avoid the kinds of compositional circularities which theories based on extensions generate. We need values which reveal how speakers can sometimes compositionally settle semantic statuses of sentences containing “true” with strictly less information about the extension of “true” than a rote application of model-theoretic tools would require.

So to construct our liberalized semantic values we need to get a better grip on which ‘restricted’ pieces of information about the set T of truths are required for various compositional processes, and when. Which restricted information we need will, of course, depends on which utterance we are considering. So we need to get a better understanding of the following utterance based notion of semantic dependence.

An utterance u compositionally semantically depends on an utterance u' if one needs information about the actual semantic status of u' to compositionally settle the semantic status of u .

Somewhat metaphorically, a compositional semantic dependence relation tells speakers confronted with an utterance when they *know enough* about which semantic properties utterances bear in a given circumstance to compositionally

settle the value of the original utterance. I make use of the terminology “semantic dependence” not only because I find it a natural label for the relation we should be out to settle, but also to signal that it is fruitfully compared to several formal and informal relations from the literature on theories of truth that tend to go by the same name. Something like my notion of semantic dependence is implicit in Kripke’s notion of *groundedness*. Versions of the concept are also treated more directly in Gaifman (1992), Maudlin (2004), and Leitgeb (2005) among others.

I bring up prior characterizations of semantic dependence for two reasons. One is to be forgiven for not providing the details of a particular formalization here, which would get us far afield. Interested readers can look to the works cited to see the variety of formal methods for dealing with semantic dependence relations. The other reason I bring up these authors, though, is by way of contrast. Plausibly there is not just one relation of semantic dependence, but many different relations which can serve different purposes. Though there are several very fruitful characterizations of dependence relations in the literature, no author has given a *compositional* semantic dependence relation of the right form for the purposes I have in mind. For example, does Sid’s (2) depend on itself?

(2) Everything I say today will be true.

On some completely legitimate ways of construing ‘semantic dependence’ the answer could be yes. (2) is clearly ‘about’ itself, for example. But (2) is *not* compositionally semantically self-dependent. In fact, part of what I’ve been trying to argue is that *no* truth-evaluable utterance can be compositionally semantically self-dependent. To be compositionally semantically self-dependent is just to involve what I earlier called a compositional circularity. So for our purposes we need to track a special, restricted relation of dependence tailored to the needs of a compositional theory—hence the new name.

All I want to do in this paper is give the *form* that such a dependence relation will have—to say what kinds of information the dependence relation will be responsive to. In particular, I want to argue that a compositional semantic dependence relation of the sort I’m after should have two special features: it should be *sensitive to extrinsic semantic information* and it should be *dynamic*. Let me explain what I mean by this.

Begin by noting a problem with what might seem like a natural form for a semantic dependence relation.

SD_1 : A compositional semantic dependence relation is a two-place relation between utterances.

On the interpretation of this claim I find troublesome, a semantic dependence relation is one which is intrinsic to the utterances compared, and could be read off of the syntactic properties of the utterances, perhaps along with information about the denotations of referring expressions or quantifier restrictors, so as to keep track of which things are ascribed semantic properties by the utterances.

Some semantic dependence relations given in the literature are broadly of this form. But the kind of semantic dependence relation we need to understand the compositional behavior of “true” requires more than this, which we can see by returning to Kripke’s classic examples. If Jones utters (J)—call this u_j —and Nixon utters (N)—call this u_n —then does u_j semantically depend on u_n , vice versa, or neither?

(J) Most of what Nixon says about Watergate is true.

(N) Everything Jones says about Watergate is false.

We’ll struggle to find an answer here because there isn’t one, unless we are told further facts. In some scenarios we need to fix the semantic status of u_j before u_n . This occurs, for example, if u_n is Nixon’s only utterance about Watergate but Jones has uttered other clear falsehoods. At other times we need to evaluate u_n before u_j . This occurs, for example, when Jones’ only utterance is u_j and the grand majority of Nixon’s utterances about Watergate are independently ascertainable as false.

Kripke raised his examples in part to show exactly this. But in the context of our discussion, the asymmetries between u_j and u_n have a special importance for how we think of our semantic dependence relation. Whether two utterances stand in that relation depends on *syntactic and semantic information extrinsic to the utterances*. To know whether u_j depends on u_n compositionally, you need to know facts about other utterances and the world first. In principle any amount of such information could be required. So we should relativize our notion of semantic dependence relation as follows.

SD_2 : A compositional semantic dependence relation is a three-place relation between pairs of utterances, and the interpretation of non-semantic vocabulary.

I’m assuming, as would be the case in many formal modeling scenarios, that the interpretation of the non-semantic fragment of one’s language contains the information necessary to fix the truth-values of sentences not containing semantic terms and (if the fragment is of interest) information about which utterances exist and their syntactic structure.

Relativizing dependence in this way allows the determination of the dependence relation in a circumstance to hang in a fairly broad sense on how things stand in that circumstance. If this were sufficient to characterize an adequate dependence relation of the sort I’m after, we might again have some useful formalizations from the literature on truth to work with. The kinds of relations most explicitly treated in Leitgeb (2005), I feel, afford us particularly good candidates. But this might still not be adequate: the semantic dependence relation hangs on more. To see this consider (9).

(9) Everything true Laura said, Tom believed.

To ascertain the truth of an utterance of (9) what information do you need to have? Information about *everything* Laura said? Again, it depends. It depends on how much of what Laura said is true, perhaps along with other facts. Consider a scenario in which we know that there is at least one truth Laura spoke—that Tom is unreliable—which Tom does not believe. It seems like at this point we know enough, in this scenario to fix the value of an utterance of (9): it is false. This means that in this scenario *relative* to the information about which utterances are true, we need fix no more truth-values of utterances to compositionally determine the status of (9).

This kind of fact should be recorded in our characterization of compositional semantic dependence because it could be relevant to how certain apparently self-dependent utterances are, from a compositional standpoint, ‘safe’. For example, for all I had said, it could have been that Laura produced an utterance of (9) in Tom’s presence. This wouldn’t affect our ability to ascertain that (9) is false in that scenario, perhaps skirting potential problems of self-reliance in Laura’s utterance of (9).

What this shows us is that the semantic dependence relation I’m out to characterize should be *dynamic*. It has the potential to shift depending on which truth-values have been allotted to utterances *using the compositional mechanisms associated with the truth predicate*. The semantic dependences of an utterance of (9) depend in a complex way on the interpretations of the non-semantic vocabulary in its quantifier restrictor and matrix, along with the semantic status of utterances we can ascertain independently of the utterance of (9) itself.

The idea here is simple and should have been clear from the moment we abandoned Truth-Extensionalism. We know, from cases like Kripke’s, that we assess utterances’ truth-values in some natural order set by the semantics of the utterances evaluated along with worldly facts. When we abandoned Truth-Extensionalism, we effectively learned that our going through this kind of procedure was not incidental—not merely a useful way for us to ascertain the truth-values utterances had for independent reasons. Rather, some such procedure, or order is responsible for *settling* the truth-values of utterances. Just as speakers naturally shift the orders in which they evaluate utterances for truth and falsity depending on non-semantic facts, they also shift them in response to the semantic facts they fix as they proceed. The final point I’ve made about the semantic dependence relation is that it should be sensitive to these kinds of shifts by dynamically interacting with the determinations of truth-values made compositionally with words like “true”.⁸ So we have the following.

SD_{final} : A compositional semantic dependence relation is a four place relation between pairs of utterances, an interpretation of non-semantic vocabulary, and a partial allotment of semantic properties.

When we relativize semantic dependence to an allotment of truth-values, it is best to think of it in a slightly different way from my first characterization of it

⁸At the end of his paper, Leitgeb (2005) does briefly consider dependence relations with this dynamic character, but does not explore their structure in detail.

on page 16. Since our new relation dynamically takes into account information about the distribution of independently fixed truth-values it may be difficult, and unnecessary, to distinguish between an utterance’s dependences on those other utterances which have, and those other utterances which have not, so far been settled in a process of allotting truth-values. Consequently it is more helpful to think of the resulting notion of semantic dependence as follows.

*An utterance u compositionally semantically depends on an utterance u' , relative to some partial allotment of semantic properties A , if a speaker whose semantic knowledge was exhausted by A might need information about the actual semantic status of u' to compositionally determine the semantic status of u .*⁹

This allows us to think of a compositional semantic dependence relation as one which gradually ‘shrinks’ as speakers acquire more information about truth-values. An utterance of “All F s are true” compositionally semantically depends (typically) on how things stand with all the F s. If we learn some of the F s are true, the utterance now semantically depends, *relative to* the information we’ve acquired, on only the F s whose values we haven’t fixed. Those are the utterances whose truth-values might still need to be fixed to compositionally settle the truth-value of “All F s are true”. As soon as we fix any F as false, the utterance *has no* semantic dependences relative to our information about semantic properties: we need no more information to compositionally settle the value of the utterance.¹⁰

The form of the dependence relation given by SD_{final} is, at last, sufficient. It *has* to be enough because we’ve exhaustively assimilated all the kinds of information into the dependence relation which are legitimate given the role it plays in constraining the compositional behavior of semantic terms. What we’ve learned is that, in general, we cannot do with anything less. As I’ve said, no philosopher I know of has explored a relation of semantic dependence which is of this particular form in any great detail. But since previous sophisticated attempts have produced relations very similar in structure, there are good grounds to think we have the tools to characterize various relations of this kind, tailored to empirical constraints and different theories of controversial uses of the truth predicate.

Characterizing a semantic-dependence relation is one key element in a compositional semantics for semantic terms like “true”. Once we have it, a natural conception of the semantic value of “true” falls directly into our laps.

Truth-Proceduralism. The semantic value of “true” is a function which maps an utterance u and any partial allotment of semantic

⁹I say ‘might’ because what might potentially be an utterance whose semantic status we need to fix to compositionally fix the value of u may turn out not to be because some *other* utterances’ values might be fixed first, and suffice.

¹⁰Note that when we shift to this understanding of a compositional semantic dependence relation, compositional semantic self-dependence need not constitute a problematic compositional circularity. It’s ‘essential’ compositional semantic self-dependence—compositional semantic self-dependence which never *disappears*—which is problematic.

properties A such that u has no semantic dependences relative to A , to a truth value.

Let $SD_{\mathcal{M},A}$ represent our semantic dependence relation as a function from utterances to their compositional semantic dependences relative to a model parameter, \mathcal{M} , and a partial assignment of truth-values to utterances, A . Then according to Truth-Proceduralism the semantic value of “true” (relative to a context, etc.) is a function F_{true} with a domain given by utterance-assignment pairs such that the utterances have no semantic dependences relative to the assignment, and model \mathcal{M} .

$$\text{dom}(F_{true}) = \{ \langle u, A \rangle : SD_{\mathcal{M},A}(u) = \emptyset \}$$

The range of F_{true} includes whatever extensional values we would like to pair with whole utterances—presumably truth-values in some set \mathcal{V} including truth and falsity.

$$\text{ran}(F_{true}) = \mathcal{V} = \{t, f, [\dots]\}$$

The “proceduralism” in “Truth-Proceduralism” is twice apt. First, it signals that the semantic value of “true” is not just an extension, or set, but a more liberalized semantic value—a special kind of procedural, or functional application. It also signals that Truth-Proceduralism takes it to be essential to the semantic value of “true” that it makes use of partial information about the set of truths in a sequential and compositional assignment of truth-values to utterances. The sequential assignment is itself a kind of procedure constitutively connected to the truth predicate’s meaning.

Truth-Proceduralism embodies the minimal departure from Truth-Extensionalism needed to overcome its challenges. We saw that Truth-Extensionalism fails because it sometimes pronounced non-defective utterances essentially semantically self-dependent: it sometimes ineliminably *required* information about the semantic status of an utterance u to compositionally settle why u had that very status. This should *never* happen if semantic properties of whole utterances are fixed by the semantic values of their parts in a non-circular way. We must be able to compositionally fix the values of these utterances without antecedent knowledge of the output of that compositional process. Otherwise a standard compositional association of u with assertoric effects is impossible, or at least pointless.

Truth-Proceduralism is the most natural way of avoiding this problem. It stipulates that the compositional effects of “true” may hang on partial information about the allotment of truth-values among utterances. It requires less than the set of truths and, by definition, exactly as much information we need about the set of truths for doing compositional work and no more. As long as a compositional semantics for “true” is *possible* it will be representable with this structure.

It is important to note that Truth-Proceduralism does not simply set the semantic value of “true” as a provisional extension in an utterance u , characterized by the values of u ’s semantic dependences. That would be to think of the

word “true” as constantly shifting in meaning in ways that would skirt the problems with Truth-Extensionalism, but also without meeting our needs for genuine reflexive applications of semantic terms. Sid would never be able to state that the conditions of his promise to speak only truths will be fulfilled, and we will never be able to talk about assertion, and its semantic and pragmatic features, in a completely general way. This is why this partial set of information about truths figuring in uses of “true” doesn’t simply act as a *pro tanto* extension, and instead serves in some other less standard compositional mode of determining compositional effects from that partial value represented by the function F over the compositional semantic dependences’ statuses and the utterance itself.

Note, of course, that Truth-Proceduralism is also only giving the *form* that a semantic value for words like “true” should take. In this way it parallels Truth-Extensionalism, which likewise tells us the shape our compositional theories will take as they integrate semantic vocabulary. Just as there are many competing ways of giving an extension for “true”, there are many competing ways of giving it a procedural value. The formulation thus leaves open, as I would like to do here, several ways of implementing Truth-Proceduralism, especially as regards utterances whose treatment is contested like paradoxical ones. But the many different ways of fleshing out this basic form to cope with controversial cases like the semantic paradoxes will all allow for conventionally acceptable ways of fixing the truth of defaulters and problematic generalities. Let me very briefly go over some options we have to do this.

Why are defaulted utterances true? Clearly because of some sort of (tacit) convention. It is a kind of convention that we plausibly *could* have had for truth tellers like (T) (especially empirical ones), but that we don’t.

(T) The second line labelled “(T)” in this paper is true.

It is not as though defaulted utterances acquire their status of being true by a mysterious non-compositional process—a fact which by happy coincidence harmonizes with its compositional make-up after the fact. Rather, speakers (at least of English) have adopted a way of using the word “true” which conventionally sets defaulters to be true (in the right circumstances). What this means is that there is a convention in the truth-predicate’s procedural value which allows utterances like Sid’s utterance of (2), in the circumstance in which it is made, to include among its compositional semantic dependences at most all the other utterances Sid made that day.

(2) Everything I say today will be true.

When enough of those dependences have “disappeared” (by having their values fixed) the semantic value of the truth predicate F_{true} is free to operate on Sid’s utterance and an allotment of truth-values to those utterances (or some subset of them), returning the value *true* if all the utterances are allotted *true*, and *false* if at least one is false. The stipulated restricted range of dependences in the compositional semantic dependence relation is what enables (2) to default.¹¹

¹¹An important question is of course when such a restriction goes into effect. In English

A somewhat more complex account will be required for semantic generalities. What should the semantic dependences of an utterance of a sweeping generality like (5) be (supposing, for illustration, that it is true)?

- (5) If a statement is true, then it's true in to predicate "true" of that statement.

Arguably *none*. There is no one actual utterance whose semantic status must be fixed before one ascertains the truth of an utterance of (5). If the generality states something true or false at all, it is able to do this through a sensitivity to allotments of truth-values which *could* obtain, as opposed to relying exclusively on the particular allotment that actually does. Moreover, not just any distributions are to be considered (e.g., those in which "true" has different meanings in English), but only those *compatible* with certain stipulations as to how the word "true" behaves.

If to fix the value of (5) we only need to appeal to hypothetical, and not actual, information about allotments of truth values, a natural way of representing the operation of "true" in settling the truth-value of an utterance of (5) would therefore be by the use of something like supervaluations.¹² We have to be careful here, though. It might be tempting to say: the compositional semantic dependences of (5), and hence its truth-value, are ascertained in part by appeal to supervaluations over semantic properties of utterances in which we hold fixed the meaning of the word "true", so that it means what "true" actually does in English. This is not acceptable. The meaning of "true" in English (and any language), I've argued, arises in part out of the compositional semantic dependence relation associated with it. If this dependence relation in turn depends on what "true" means in English, we won't have picked out a class of truth-value distributions over which to perform supervaluations at all. (This is, in effect, just another manifestation of the special kind of relationship (5) has to itself that led to compositional circularities in conjunction with Truth-Extensionalism.) So instead of ascertaining dependences, or truth-values, by reference to what "true" means, our dependence relation must stipulate those features of the use of "true" which should be constitutive of its use, and hence constrain the supervaluations used to assess semantic dependences, and truth-values. These will presumably include stipulations for how "true" should interact with logical connectives, and with itself. The latter conventions will help settle that (5) has a stable truth-value over the entire range of supervaluations, showing it to have no semantic dependences. Accordingly, we have an explanation of how (5) avoids generating compositional circularities. Similar strategies will help with other semantic generalities.

the restriction seems to occur only when the utterance carries information about the truth of at least some other utterances which do not 'simultaneously' need to default. I won't be exploring the details of defaulting conditions here.

¹²This is especially welcome since some of the most sophisticated relations of semantic dependence, like Leitgeb's, already appeal to supervaluations. Other methods, like taking the consequences of deductive rules constitutive of the use of "true" might also do this work, though I am suspicious of the empirical viability of such methods.

What I'm giving here is, of course, only the barest sketch of how we might proceed, and is only one of many options. The point is that Truth-Proceduralism opens up new opportunities for understanding the compositional behavior of semantic generalities which were closed off by Truth-Extensionalism. The behavior and logical properties of these various options are worth developing and exploring with many familiar techniques we have on hand, like supervaluations or use of deductive methods. Any such technique will be permissible in use with Truth-Proceduralism as long as it does not presuppose the availability of an utterance u 's actual truth value in the compositional processes which determine that status.¹³

The details of any such account are less important for my purposes here than the possibility of some such account—a possibility which does not exist for the Truth-Extensionalist. This possibility is afforded through the interaction of two key elements: the dynamic semantic dependence relation for utterances containing “true”, and the rule for assigning truth-values along those dependences. The dependence relation and the rule for assigning truth-values interact with each other in a way that makes both integral to understanding the meaning of a word like “true”. The dependence relation hangs, as we have seen, on the provisional outputs of the rule. This is what motivated our final adjustment to the form of the semantic dependence relation above in SD_{final} . The rule associated with “true”, however, proceeds as a function of how things stand with the semantic dependence relation. That relation helps settle when compositional processes can, with the help of worldly information, fully engage.

4 Context, Expression, Logic, and Paradox

Now that we have our schematic form for the operation of semantic predicate denotations, I want to turn to consider some implications of adopting them. In keeping with my main goal in this paper, I've been abstracting away from concerns raised by the semantic paradoxes and focusing on the independent question of how to construct a compositional semantics for truth. My contention was that if we started with this seemingly simpler problem, we might arrive at conclusions about truth which have ‘downstream’ effects—new implications for theories of language and logic which may in turn give us new tools for approaching the core problems which have mainly preoccupied contemporary theorists of truth. In this last section, I want to very briefly show that several important theoretical domains are indirectly affected by the shift to procedural semantic values. Let me begin with a surprising case: treatments of context-sensitivity.

In §2, I noted that the problem with using standard model-theoretic techniques in explaining productive uses of “true” had nothing to do with the particular distributions of truth-values they gave rise to. It was the methods which the

¹³Or more carefully: it does not presuppose that availability in any such non-defective utterance. The idea that procedural values can fail to engage has obvious valuable applications in explaining defective utterances containing semantic vocabulary, like truth-tellers for example.

theories used to reveal how composite wholes inherited their truth-values from the meanings of their parts. Accordingly, we made the minimal modifications to semantic values required to skirt this problem. But this leaves open an important question: once we accommodate these new kinds of semantic values, is there room for the truth-value distributions across utterances to change as well? That is, are there truth-value distributions to whole utterances in procedural systems which won't arise in the model theoretic setting?

In at least one important sense, the answer is “yes”. The liberalization of semantic values leads to a commensurate liberalization in the potential allotments of truth-values to whole utterances. The reason is that procedural semantic values give us new ways to allot truth-values to tokens of the same semantic type—they yield a new form of ‘token-sensitivity’ which is unlike familiar forms of context-sensitivity. I’m not going to argue here that any of these more liberalized distributions arise in any actual natural language, but I want to explain why the form of token-sensitivity that arises from procedural values is new, and of potential theoretical interest.

On a familiar, broadly Kaplanian representation of context’s involvement in the semantics of predicates, predicates are associated not simply with an extension, but a function from contexts to extensions. When an utterance is produced in assertion, the context of utterance interacts with the standing meaning of constituent predicates to generate extensions which can then figure in the compositional determination of assertoric effects in a familiar fashion. In this way tokens of the same sentence type may be paired with different truth-values. These methods might be used to show how my utterance of “Jones is tall” could be used to say something different from yours. This might be because, say, we are asserting Jones is tall relative to different standards of height. So even though the utterances have the exact same surface syntactic structure, and parts with the exact same standing meanings, the involvement of context may generate a kind of what I’m calling token-sensitivity—a distinction between the assertoric content paired with distinct tokens of the same sentence type.

Some authors have, on grounds having to do with paradox, supposed that uses of semantic predicates exhibit some kind of token sensitivity. Gaifman (1992) has pushed for such a treatment of what he calls the ‘two line paradox’:

line 1: What’s written on line 1 is not true.

line 2: What’s written on line 1 is not true.

What’s written on line 1 is paradoxical for familiar reasons. One might wish to call it untrue on those grounds. This is precisely what happens in line 2. If we can consistently pronounce paradoxical utterances untrue, one might expect that we are able to do so because new tokens of the same type are somehow immune to the defects of the original paradoxical utterances. Glanzberg (2001) sees a similar difference between the different tokenings of sentence types containing “true”, and eventually posits a special kind of context sensitivity in quantifier domains to account for this fact.

Again, the issue of whether or not utterances containing “true” exhibit a form of token-sensitivity, as these authors argue, goes beyond the scope of this paper. What is surprising about procedural values is that they give us an entirely new potential locus for some such token-sensitivity, one that differs from standard treatments of sensitivities in predicate denotations and quantifier domains to features of context. This is because procedural semantic values have the potential to distinguish between sentence tokens on the basis of their relationships in a hierarchy of semantic dependences *without* any influence on the standing meanings of the constituents in those tokens *or even* the meanings of those constituents ‘in context’. This, in turn, is because a compositional semantic dependence relation which helps constitute the procedural semantic value is itself token-sensitive: different tokens of the same sentence type may have different compositional semantic dependences.

To understand what I mean by all this it might be useful to look at our increasingly familiar, though helpfully simplistic, case of defaulting. Consider a more stilted sentence that Sid might have uttered.

(11) Everything Sid says on March 15th 2011 is true.

Suppose Sid utters (11) on March 15th 2011, a day on which he additionally produces only some small number of uncontroversially true statements. Then speakers will count this utterance, u_s , true in the now familiar manner. Suppose, at some point, Ira shares in Sid’s assessment and utters (11) as well. Again, if the circumstances are as I mentioned, this new utterance u_i will count as true.

Sid’s utterance u_s and Ira’s u_i have the same words, with the same arrangement, arguably with the same meanings, even in context. They also have the same truth values. What is interesting about them is that, from a compositional perspective, they are true *for different reasons*. Sid’s utterance requires a defaulting mechanism to engage. Were it not for the conventions of English, Sid’s utterance might have been defective. But Ira’s utterance requires no such mechanism. It does not ascribe a truth-value to itself in the way that Sid’s utterance does in the circumstance I have described, and because of this, it can be compositionally sensitive to a broader distribution of truth-values than Sid’s utterance. Sid’s utterance u_s *must* not require information about itself while its truth-value is being settled—it must not be compositionally semantically self-dependent. However Ira’s utterance may, should, and seemingly *is* compositionally responsive to that very same utterance u_s (just consider if Sid had said different things at that moment). So the operation of the truth-predicate’s semantic value in these two utterances is different—it responds to different bodies of semantic information, and does so in different ways.

Where is the locus for this difference in compositional behavior? Not in the standing meanings of the expressions used in u_s and u_i . Nor is it in any standard Kaplanian interaction with context of utterance. We have not given any room for such interactions in setting up procedural semantic values. Nor, again, does it lie in a form of context-sensitivity in quantifier domains. Moreover, it’s not clear that we should *want* to shift to these forms of context dependence to

account for the differences between u_s and u_i . Procedural semantic values have the possibility of the compositional sensitivity to sentence tokenings built into their operation. And trying to remodel these distinctions in u_s and u_i with more familiar kinds of context-sensitivity lands us back in a misleading form of Truth-Extensionalism. Doing this will, for example, eradicate the possibility of genuine reflexivity in utterances like u_s .

So a compositional asymmetry between u_s and u_i arises even though there is *no* difference in their syntactic *or* semantic make up. They involve constituents with the same semantic values (even in context), with predicates designed to record the same properties, and all their constituent values are combined in the same way. Nonetheless they exhibit different compositional operations. There is no great mystery about how this occurs. It simply follows from the new kind of entities that procedural semantic values are.

Now, the differences between the compositional workings of u_i and u_s might seem insignificant because they don't result in a difference in truth-value. But what examining this case shows is that they *could have*. There is 'room' between u_s and u_i for a different allotment of truth values even though, for obvious reasons, speakers don't want to situate one there with stipulations on how "true" should be used. But the same compositional 'distance' exists for other distinct utterances of tokens of the same type which are hierarchically arranged in terms of their semantic dependences. This notably occurs for the utterances in the aforementioned two-line paradox.

It is a difficult question whether speakers do or should make use of the token-sensitivity afforded by the shift to procedural semantic values. What we've learned is that English is built up using a *kind* of semantic value which structurally allows for entirely new ways that utterances' truth-values depends on more than the semantic properties of the sentence types uttered.

Now that we appreciate how the shift to procedural semantic values can affect our understanding of context-sensitivity, we are in a position to appreciate its potential to affect a separate domain of key philosophical interest: logic. Earlier in this section, I emphasized that the move to liberalize semantic values wasn't motivated by the need to accommodate new patterns of truth-value allotments. What we've now seen is that is that liberalizing our semantic values in the weakest way possible ends up allowing new possible kinds of truth-value allotments nonetheless. Consequently, even if speakers of no actual natural language take advantage of these liberalized allotments, systems with procedural semantic values have a different logic than systems without them. What these logics look like depends on hard questions I won't try to answer here—for example, whether we should treat special features of our procedural values such as the semantic dependence relations they advert to as somehow elements of logical forms and, if so, how. Regardless of how we settle those questions, it behooves us to investigate these logics, and their foundations, if we want to fully understand the 'logic of truth'.

Lastly, and most briefly, I want to note that the shift to procedural values has the potential to affect broad questions in the foundations of linguistic expression. In §2, I tried to stress that although we should be wary of associating a semantic

predicate like “true” with an extension in a compositional theory, the reasons for caution did not speak against there being a property of *truth*, in my minimal sense of “property”—a set of truths. But even if we can preserve our ability to think about truth as a property, we are forced to confront new ways for thinking about what it is for a predicate to *express* this property. When we assign a predicate an extension or intension as its semantic value, the question of how the predicate represents, or expresses, the property is answered trivially. With procedural semantic values the question requires a non-trivial answer because of the new need to effect a strong separation between the semantic property on the one hand and the compositional tools we have to record it on the other.

Whatever answers we give to questions about how semantic predicates represent or express properties, one thing is sure: these predicates represent, or express, those properties in a non-standard way. If we are forced to reconceive how it is that a predicate may adequately represent a property, we are thereby forced to reconceive *what it takes* for a property to be adequately represented by a predicate. This, in turn, may teach us new things about expressive power. Once we change our perspective on what is required of a language for it to count as containing a term expressing a semantic property, we change our perspective on one of the issues at the forefront of current research on truth: can a language contain the means for expressing *all* semantic concepts, or all suitably interesting semantic concepts? And if not, what are the limits on a language’s semantic expressive power?

I say all this to emphasize that the shift to procedural semantic values has a rather surprising theoretical reach. It may seem that simply shifting between two different ways of associating whole utterances with semantic properties, as opposed to shifting between two distributions of semantic properties themselves, is unlikely to effect any important differences outside the domain of compositional semantics. But we’ve now actually seen three domains of inquiry—linguistic context-sensitivity, logic, and linguistic expressive power—which are substantially affected by the transition to procedural values.

These are a few of the ‘downstream’ consequences that I alluded to earlier. Though tracing out these consequences relies in no way on any controversial view of semantic paradox, as I’ve been at pains to stress, it’s clear that these consequences also have great potential to teach us about paradox. After all, the areas of logic and expressive power are at the very center of debates concerning the adequacy of various formal theories of truth. But procedural semantic values should likely also influence discussions of paradox in a much more direct way.

We’ve just seen that semantic predicates, for reasons having nothing to do with paradox, should be outfitted with special kinds of semantic values unlike any other predicates in our language. Clearly if it turns out that these predicates also exhibit special behavior in other, more controversial uses, it should look tempting to tie the peculiarities of those controversial uses to the procedural semantic values. It would be the first thing we would have checked had we been giving a compositional semantics for “true” before working on the paradoxes. And if a theory successfully make such a link it would have a powerful motivation for a treatment of the paradoxes, which might threaten to give rival views an

ad hoc character. I'll unfortunately have to leave the question of whether any existing theories have the potential to forge the connections required here for another occasion.

A final caveat: All these ideas only apply if we *start* from the idea that “true” should admit of a compositional semantics like any other term, and work from that starting point gradually into the more treacherous territory of paradox. It may be, as I conceded at the outset of the paper, that those excursions will lead to a kind of *reductio*, forcing us to abandon the apparently modest starting assumptions of this paper. I acknowledge that we cannot tell before investigation this will not be the case. The final lesson we can draw from my work here concerns this possibility and what impact this would have due to the connections between compositional and foundational semantics I traced out in §1. Theorists of truth proposing the relevant revisionary measures need to be more forthcoming about the effects this result has both on foundational work in the philosophy of language, and the formalisms which were once undergirded by that foundational work. Abandoning the foundational goals set at the outset of this paper threatens the very point of thinking about truth, at least in the kinds of sophisticated technical setting in which truth-theorists are accustomed. The formalisms we use to reflect about truth have always owed their significance to the ties—potential or actual—that they have to the intentional properties of the language we use, or the norms governing linguistic use. When such ties are severed, or perturbed in the way that some have proposed, not only will the grand majority of our projects in foundational semantics founder, but the fate of formalisms relying on them will need new foundations of another sort, lest they become *mere* technical apparatus. As to what the nature of these new foundations could be, I find it difficult to speculate.

References

- J. Beall (2009). *Spandrels of Truth*. Oxford University Press.
- H. Field (2008). *Saving Truth From Paradox*. Oxford University Press.
- H. Gaifman (1992). ‘Pointers to Truth’. *The Journal of Philosophy* **89**(5):223–261.
- H. Gaifman (2000). ‘Pointers to Propositions’. *Circularity, Definition and Truth* pp. 79–121.
- 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.
- A. Gupta & N. Belnap (1993). *The Revision Theory of Truth*. MIT Press, Cambridge, MA.

- S. Kripke (1975). 'Outline of a Theory of Truth'. *Journal of Philosophy* **72**(19):690–716.
- P. Lasersohn (2006). 'Compositional Interpretation in Which the Meaning of Complex Expressions are not Computable from the Meanings of their Parts'. *Manuscript* .
- H. Leitgeb (2005). 'What Truth Depends on'. *Journal of Philosophical Logic* **34**(2):155–192.
- T. Maudlin (2004). *Truth and Paradox: Solving the Riddles*. Oxford University Press, Oxford.
- V. McGee (1991). *Truth Vagueness and Paradox: An Essay on the Logic of Truth*. Hackett Publishing, Indianapolis, IN.
- G. Priest (2006). *In contradiction*. Oxford University Press.
- A. Tarski (1956). 'The concept of truth in formalized languages'. *Logic, semantics, metamathematics* pp. 152–278.