Metasemantics and Contrastivity

Here lies the record of a metaphysical rabbit hole that I fell into sometime during the Spring 2021. I'm not sure if anything will become of this. I found the questions to be intriguing and (I still think) important, but without any real results, there's not much that I can do with it.

1. Basic metasemantics

Metasemantics is the branch of metaphysics that is specifically concerned with the metaphysics of meaning. Thusly stated, there is plenty of room for debate over the precise explanatory target of metasemantics; however, for the purposes of this essay, I will make a few declarations by fiat to narrow down its scope, in the interest of making things more manageable.

Firstly, metasemantics aims to explain the facts of meaning, but there are different understandings of what those facts are. According to the traditional approach—inspired by Frege, Russell, Tarski and others—meaning is understood in truth conditional terms. The meaning of a sentence is then identified with its truth condition, the meaning of a name is identified with its referent, and the meaning of other expressions is understood as their contributions to truth conditions. However, there are other approaches. Some views cast meanings as conceptual roles, patterns of usage, or norms of usage. Still, other views cast meanings as concepts, or some such entity that is more fine-grained than reference. Finally, other views understand meanings as the dynamic contributions made to conversations.

My first stipulation is that I will understand the facts of meaning along traditional lines. Thus understood, the explandanda of metasemantics are such facts as *that t refers to o*, *that F applies to x,* and *that S is true iff p,* etc.

Although we say that metasemantics is interested in explaining such facts, there are several different readings of what kind of explanation we may be after. For instance, one may be interested in the essential natures of the semantic components of the semantic facts. This interest would give rise to an inquiry into the *nature of reference* and the *nature of truth*—both inquiries that have their own long and tortuous histories. However, neither of these will be the chief purpose of metasemantics as I will delineate it. Instead, I will stipulate that metasemantics is chiefly interested in the *grounds* or *determinants* or *underlying states that fix* the facts of semantics. Within the purview of this explanatory endeavor, the nature of the semantic qualities, such as reference and truth *in themselves*, is not the primary phenomenon to be interrogated. Rather, the aim is to uncover the states, facts, or features that *determine* the facts of reference, truth, etc.

Once metasemantics is understood in this way, we can widdle down it's target by appealing to another well-established semantic phenomenon: composition. Plausibly, the semantic features of complex expressions—including sentences—are determined by the semantic features of their simpler components, through semantic means. Thus the immediate determinants of the complex facts are within the purview of semantics, not metasemantics. On the other hand, the facts that require a distinctly *metasemantic* inquiry are the ones that pertain to simple expressions.

With these three stipulations, we thus arrive at the questions that I take to be distinctly metasemantic. Let *t* be a term, *F* a predicate, and *o* and *x* be objects. The representative metasemantic questions are:

- What non-semantic facts ground the fact that *t refers to o*?

- What non-semantic facts ground the fact that *F applies to x*?

2. Grounding as contrastive

Since the metasemantic questions under consideration are essentially grounding questions, it's important to keep some general features of grounding in view.

There is a long-standing debate over the proper relata of the grounding relation. Some say it is objects and others say it is facts. Given that our ultimate target is the semantic *facts,* we had better just stick to facts. Thus understood, the grounding relation has the structure *that ϕ grounds that ψ*.

However, Schaffer (2012) argues that this isn't the full explicit structure of a grounding relation. Instead, he argues that we must understand grounding as implicitly contrastive. Thus, in his view, the *real* relata of a grounding relation are not plain facts, but rather *differences*. Consequently, the real structure of a grounding relation is rather:

(1) The fact that φ rather than φ* grounds the fact that ψ rather than ψ*

Here, φ and ψ are obtaining facts, and φ* and ψ* are non-actual alternative facts that are ruled out. So, for example, on Schaffer's view, merely saying that *Tom's c-fiber being stimulated grounds Tom feeling pain* is insufficiently explicit. Instead, a proper grounding claim would be *that Tom's c-fibers being stimulated rather than not stimulated grounds the fact that Tom feels pain rather than not feeling pain.*

Shaffer gives several reasons for thinking that this is the proper structure of grounding claims. For one, grounding is thought to be analogous to causation, and as Schaffer observes, it is widely conceded that causation has a contrastive structure. This is evidenced by the fact that causal claims can change truth value depending on where we put the emphasis. Socrates drinking *Hemlock* at dusk caused him to die. *True.* But it is not true to say that Socrates drinking Hemlock *at dusk* caused him to die. To explain this divergent behaviour, observe that a difference in emphasis suggests a difference in contrast. For another reason, grounding is taken to support (or be a species of) explanation. And it is also widely recognized that explanation has a contrastive character. To borrow an example from van Fraassen, explaining why Adam ate *the apple* is a different task from explaining why *Adam* ate the apple or why Adam *ate* the apple. Finally, Schaffer's main argument takes off from the widely-accepted claim that grounding is transitive. But, as Schaffer argues, if we ignore the hidden contrastive elements in the grounding relation, then we can produce apparent counterexamples to grounding's transitivity. I won't rehearse the counterexamples here, but they are interesting in their own right.

3. Metasemantics as contrastive

Suppose that we adopt the claims of the two preceding sections: *metasemantic explanation is a species of grounding* and *grounding is inherently contrastive*. We then get the result that *metasemantic explanation is inherently contrastive*. Let's explore this idea a bit further.

Consider the case of singular reference. Ordinarily, one would think that the grounds for a fact of reference consist in some non-semantic relation born between the referring expression and its referent. The (non-contrastive) form of such a metasemantic explanation is:

(2) That *t* bears *r* to *o* explains that *t refers to o*.

(*t* is a referring token, *r* is a non-semantic relation*, o* is the referent.) But if grounding is really contrastive, then we need to complicate the structure of metasemantic explanation by including the contrastive elements. The *real* structure of metasemantic explanation would then be:

(3) (1) *t* refers to *o* rather than *c* because there is a relation *r* such that *t* bears *r* to *o* and *t* does not bear *r* to *o*.

The contrasted object *c* is an alternative object that we are endeavouring to rule out as the referent. The relation *r* is some non-semantic relation born between *t* and *o* and not born between *t* and *c* that purports to explain the difference. When a non-semantic relation succeeds in this role, we say that it "supports" the explanation.

Let's give an example. Suppose that we are walking on a beach, and we see a dim light in the distance. I point to it and say "that's a lighthouse." We now ask: why does my token "that" refer (on this occasion) to the object—call it *X*—that is the cause of the light, rather than, say, the mountains on the distant shoreline? Presumably, this has an easy answer: it is because I was intending and gesturing to bring your attention towards *X*, which I was also perceptually focused on, and not towards the mountains on the distant shortline. If we really wanted to get into the details, we can further explain why my intention and perceptual state is about *X* rather than the distant mountains, by appealing to the mechanisms that underwrite my perceptual focus and the causal relations between *X* and my perceptual state.

In offering this answer, notice that I am not thereby attempting to say what reference *is*, or to state *general* conditions for the fixation of reference, or even to explain why my term refers to what it does, as opposed to *anything else*. I am merely bringing up salient facts about my speech in order to explain the given contrast.

In several other places, I have called this sort of explanation a "selective explanation of reference." This is to highlight the fact that the point is to explain why the particular referent was *selected* at the exclusion of other objects that aren't actually the referent. I have also claimed that this kind of metasemantic explanation is not particularly demanding. It usually isn't too hard to explain why, say, "that" (as in "that's a lighthouse") refers to one object X (the cause of the light) rather than the distant mountains.

What sort of relations explain selective explanation? Intuitively, there may be a plurality of factors that are relevant to reference selection. We may want to appeal to causal relations, relations of speaker deference, societal linguistic conventions, perceptual relations, cognitive-computational relations, nomic regularities, and so on. Generally speaking, I think that the relations that support selective explanation have to do with the production of the referring item, but I leave it fairly open-ended as to their specific natures. (One caveat: these relations have to be *genuinely explanatory*, so they must be, in some robust sense, *real*.)

4. Two theses on metasemantics understood contrastively

It's quite natural to think that the semantic facts are (in fact) grounded in the non-semantic facts. That is just to say that they aren't *brute* or *fundamental*. At any rate, I won't discuss the potential

counterexamples here (c.f. Brandom (1996)). I will take it as given that the semantic facts are all grounded.

If we combine this thought with the *contrastivity* of grounding, then questions arise as to how exactly this works. There are at least two theses, of different strengths, that outline how the selective explananada receive their answers. The first one that I will advance is this:

Contrastive Groundedness Thesis (CGT): For all terms *t*, for all objects *o* and *c*, *t* refers to *o* rather than *c* if and only if there is a non-semantic *r* such that (i) *t* bears *r* to *o*, and (ii) *t* does not bear *r* to *c*, and (iii) *t* bearing *r* to *o* rather than *c* explains why *t* refers to *o* rather than c.

Briefly put, this thesis says that for any contrast, the request for a selective explanation will always have an answer. There will always be *some* non-semantic relation that can support the explanation.

The contrastive groundedness thesis represents a fairly weak thesis about how the reference facts are grounded. It does not require that all reference facts, for all of the various terms that we use, have similar grounds. Indeed, it does not even require, for a single given term and referent, that all *contrasts* have the same grounds. All it says, of the fact that *t refers to o*, is that *for each contrasted object c, there will be some r that explains why t refers to bo rather than c.* This is consistent with the grounding relation differing for different contrasted objects.

Now let's compare CGT with a stronger thesis about how the reference facts might be grounded. If we move one step above CGT, we obtain this thesis:

The Singular Grounding Thesis (SGT). For all terms *t*, for all objects *o*, *t* refers to *o* if, and only if, there is a nonsemantic *r* such that (i) *t* bears *r* to *o uniquely,* and (ii) for all other objects $c \neq o$, *t* bearing *r* to *o* rather than *c* explains why *t* refers to *o* rather than *c*.[1]

This thesis is stronger precisely because it requires, for each referring term, that there be only *one* non-semantic relation that supports all selective explanations of reference. (Remember, we're considering only the relations that have real explanatory power. Since the relations that we're considering are sparse, this thesis isn't trivial.)

We can think of CGT and SGT as representing two different *degrees* or *grades* of inflationism in the theory of reference. A view that subscribes to CGT alone (and rejects SGT) is relatively less inflationary because it imposes fewer demands on our metaphysics of reference. On the other hand, a view that subscribes to SGT imposes the stronger demand that each instance of reference is susceptible to a uniform explanation that explains all contrasts; hence it is more inflationary. Nonetheless, both of these views count as versions of inflationism (as opposed to deflationism) because they claim that the facts of reference are *grounded* in the non-semantic facts, so they deny the core deflationary thesis that the reference facts are explicable solely in terms of the trivial schema *"A" refers to A.*

Now that we have these two theses on the board, we're finally in a position to ask the question that is the main topic of this discussion: supposing that either CGT or SGT is true, *which one is it?* Should we subscribe to the moderate thesis, SGT, and hold that only one relation suffices to explain the selection of reference per term? Or should we prefer the weaker thesis, CGT, and allow for a plurality of

---

[1] The uniqueness clause in (i) assumes that *t* determinately refers to only one object. This is an idealization. To account for the phenomena of semantic indeterminacy, adjustments would have to be made, which I will not consider here.

explanations for each term, depending on which contrast is at issue? (I have never before seen this exact question presented in the literature. It seems to me to be profoundly important. However, as we'll see shortly, it can quickly become messy.)

5. Arguments in favour of a CGT-only view?

To argue for a CGT-only view one must find reason to think that metasemantic explanation is inherently contrast-relative. In other words, different contrasted objects call for different explanations, even for the same term and referent.

  Presumably, the most direct way to argue for such a view is by finding a counterexample to SGT. I have given this quite a bit of thought, and I've come to the conclusion that it is quite demanding; there are a lot of moving parts and each one of them has to be convincing on their own. To be specific, a counterexample to SGT must be such that there's a term, t, objects $o, a,$ and $b,$ and: (i) t refers to o, (ii) t does not refer to a or b ($a \neq b$), (iii) that t refers to o rather than a is explained by the fact that t bears r_1 to o rather than a, (iv) that t refers to o rather than b is explained by the fact that t bears r_2 to o rather than b, (v) $r_1 \neq r_2$, (vi) that t refers to o rather than a cannot be explained in terms of r_2, (vi) that t refers to o rather than b cannot be explained in terms of r_1, and finally (vii) that t refers to o rather than a and that t refers to o rather than b cannot both be explained by another relation r that t bears to o. Not only that, but t must refer to o unambiguously throughout the example; the variation behaviour cannot be chalked up to semantic ambiguity.

  I will return to this topic if I ever think of something good. Given the above description, one thing must be clear: the viability of such a counterexample will crucially depend on the restrictions imposed on the relations r_1 and r_2. They must be distinct (as per ii). They also cannot be factors of a common relation r that can explain reference relative to both contrasts (as per vii).

  Evidently, the difference between CGT and SGT really hinges upon the *individuation* of the relations that are genuinely *explanatory*. That's unfortunate, because those sorts of questions are very sensitive to metaphysical intuitions and give rise to irresolvable disagreements and aren't particularly tractable. (Who's to say where to draw the line between the sparse properties and the plenitudinous ones?)

6. An argument in favour of an SGT-view?

When I first discovered this question, I thought it was possible to prove SGT from CGT by employing some extra metaphysical assumptions. In hindsight, the exercise was self-indulgent, because without solving the "individuation of relations" problem, it's unclear what this accomplishes.

  Regardless, here is what I wrote [what follows is an older piece of writing]:

First, we need a metaphysical assumption (MA) about the compositions of relations. Basically, we need for it to be the case that we can conjoin any two relations to form a new relation. The class of relations cannot be so sparse that the conjunction of two relations fails to exist. Here is the assumption made explicit:

(MA) If x bears r_1 to y and x bears r_2 to y, then there is an r_3 such that $r_3 = r_1 \wedge r_2,$ and x bears r_3 to y.

Remember, we're thinking of these relations as *genuinely explanatory* (dare I say, real) fine-grained

entities—not just sets of ordered pairs. As such, there is no *logical* guarantee that the conjunction of two relations will always yield an existing relation. This assumption belongs to metaphysics, not to logic.

The other assumption we need is an *assumption about contrastive explanations*—let's call it *ACE*. Roughly, we need for it to be the case that contrastive explanations can be combined using the truth-functional connectives. Formally speaking, we need for it to be the case that:

(ACE) If $<A_1$ *rather than* $B_1>$ *explains* $<X$ *rather than* $Y>$ and $<A_2$ *rather than* $B_2>$ *explains* $<X$ *rather than* $Z>$ then $<A_1 \& A_2$ *rather than* B_1-*or*-$B_2>$ *explains* $<X$ *rather than* Y-*or*-$Z>$.

Is this principle true? It certainly seems to be true. Consider this piece of reasoning:

(-) <I wanted to drink something caffeinated rather than non-caffeinated> explains <I chose to drink coffee rather than hot chocolate>
(--) <I wanted to drink something hot rather than cold> explains <I chose to drink coffee rather than Red Bull>
Therefore:
(---) <I wanted to drink something caffeinated and hot rather than non-caffeinated or cold> explains <I chose to drink coffee rather than hot chocolate or Red Bull>.

This inference strikes me as clearly valid. But in order for it to be valid, it relies on the principle ACE.

There is one final assumption to make for the purpose of this proof. However, note that this last assumption will only be made for the sake of taming the argument; we'll consider relaxing it afterwards. This is the assumption that the universe is finite. We will prove the theorem under this condition, and then afterwards look at the prospects for relaxing it.

Theorem: If the universe is finite, and the metaphysical assumption (MA) and the assumption about contrastive explanation (AE) each hold, then the CGT is equivalent to SGT.

Proof. That SGT entails CGT is the trivial direction of this proof. Just take the *r* that explains reference no matter the context, and it will (*a fortiori*) explain reference relative to each contrast class.

The difficult direction is proving that CGT entails SGT. So let's assume CGT, and let *t* be a term that refers to *o,* and let U^- be the universe of objects with *o* removed (i.e. all of the objects distinct from *o*). (Remember, we're assuming that U^- is finite.) The burden is to prove that there is an *r* such that, (i) *t* bears *r* to *o*, (ii) *t* does not bear *r* to anything else, and (iii) for all $c \in U^-$, *r* supports a selective explanation of reference with respect to *c*. (Notice that we do not need to do any further work to prove the other direction of the biconditional embedded in SGT, since the last clause of the right-hand-side implies that *t* refers to *o*. It says there's an explanation for *t* referring to *o,* and "x explains why y…" is factive.)

Here is the proof strategy. The contrastive groundedness thesis implies that for each *c* in U^-, there will be some relation r^c, such that $<t$ bears r^c to *o* rather than $c>$ explains $<t$ refers to *o* rather than $c>$. If we take these relations and *conjoin them,* we can thereby form a new relation that can support selective explanations when the contrast class is all of U^-.

First, number the $c_i \in U^-$ from 1… n. Then, the pairwise explanation conjecture implies that there will be $r^{c1}, r^{c2}, \dots r^{cn}$ such that:

1. $\langle t$ bears r^{c1} to o rather than $c_1 \rangle$ explains $\langle t$ refers to o rather than $c_1 \rangle$
2. $\langle t$ bears r^{c2} to o rather than $c_2 \rangle$ explains $\langle t$ refers to o rather than $c_2 \rangle$
...etc.

Now define $r^{c1+c2} = r^{c1} \wedge r^{c2}$ (t bears $r^{c1} \wedge r^{c2}$ to $o \equiv_{df} t$ bears r^{c1} to o & t bears r^{c2} to o). MA then implies that t bears r^{c1+c2} to o. Now, given ACE, it follows that r^{c1+c2} supports the following contrastive explanation:

(1+2) $\langle t$ bears r^{c1+c2} to o rather than any $c \in \{c_1, c_2\} \rangle$ explains $\langle t$ refers to o rather than any $c \in \{c_1, c_2\} \rangle$.

Now, given MA and ACE, we can iterate the process of conjoining relations r^{c1}, r^{c2}, ... r^{cn} to yield selective explanations relative to contrast classes that include c_1, c_2, ... c_n. Since the domain was assumed to be finite, this process will terminate with a relation r that supports contrastive explanations of reference relative to all of U^-. If r supports such explanations relative to all of U^-, then it will also support such explanations for any subset of U^-.

We have now proved the third part of our desired claim. The only remaining part to prove is that t bears r to o and *only o*. That t bears r to o is guaranteed by the construction of r and MA, so there's no work to be done there. As for the claim that t bears r to o *only,* this can be proved by reductio. Suppose that there's an $o^* \neq o$ such that t bears r to o and t bears r to o^*. By assumption, t refers to o; and t does not refer to o^*. Thus, by the CGT, there was some r^{o^*} that relates t to o but does not relate t to o^*, and this r^{o^*} was part of the construction of—and hence one of the conjuncts of—the relation r. Hence t does not bear r to o^*—a contradiction. Therefore, t bears r to o uniquely. $\square$

It is worth putting all of this into bite-sized form. What is the upshot of this proof? Well, we started with the observation that *for a particular contrast,* it does not seem too difficult to offer contrastive explanations of reference. But that observation alone allows for these explanations to vary depending on which objects are contrasted. What this theorem shows is that, under certain assumptions, we can compile the contrast-dependent explanations into a single contrast-invariant explanation. We thus forge together a word-world relation that can explain the selection of reference relative to any contrast class whatsoever.

Notice that defining the relation r, which supports selective explanations relative to any contrast, is simply a matter of compiling whatever information is relevant to offering an explanation *relative to a given contrast.* The theorem thus allows for the contrast-invariant explanations of reference to be highly complicated, as they would be if r is the conjunction of many non-repeating elements. In that case, r may not admit of any simple or natural description. But for our purposes, that's *fine.* It's enough that such an r will always exist for each referring term.

This theorem also has the virtue of revealing what exactly needs to be assumed in order to guarantee the existence of contrast-invariant explanations of reference. Specifically, we needed to assume (I) the contrastive groundedness thesis, (II) the metaphysical assumption (MA), (III) the assumption about contrastive explanations (ACE), and (IV) that the domain of objects is finite.

Of all of these assumptions, it is fair to say that the last one is the weakest link. A theory of reference shouldn't assume that there are only finitely many things in the universe. But as I have said, this was only a simplifying assumption; it is not an outright commitment of the theory. So now that we have climbed the ladder utilizing this assumption, it's time to see how we might throw it away.

One option is that we could try to run an analogous proof without the assumption of finitude. To

pursue this strategy, we would need to introduce strengthened versions of MA and ACE that handle infinite conjunctions. We would then have to face the question of whether the infinitary extensions of MA and ACE are true. And in order to properly address these questions, we would have to rely on some fairly powerful logical resources and a great deal of intuition. It seems to me that going this route could turn into quite a distraction. For this reason, I will not pursue *formally* proving the uniform explanation thesis for infinite domains.

Rather than providing a formal proof, I will instead offer a justification that relies on an informal observation. It appears that, in practice, when giving contrastive explanations of reference relative to various contrast cases, we will probably end up repeating ourselves over and over again. *Why does "that" refer to X rather than the mountain standing to its left?* Because "that" bears such-and-such relation (mentioning direction of focus) to *X* rather than the other mountain. *Why does "that" refer to X rather than another mountain standing to its right?* Because "that" bears such-and-such relation (*again,* mentioning direction of focus) to *X* rather than this third mountain. And so on. Even if there were an infinite number of differently-located objects to consider as contrast cases, the explanation will always be the same.[2] Therefore, we shouldn't need to continually conjure up new relations between "that" and *X* in order to explain this instance of reference relative to novel contrast cases that are differently located in space. We don't need new relations to explain *every* difference; at most we only need new relations to explain each *kind* of difference (however it is that differences get divided up into kinds in this context). If this is right, then we can justifiably expect that there won't be a large number of non-repeating conjuncts in the contrast-invariant, explanation-supporting relation *r*. I thus expect that, for any term, the non-repeating conjuncts will be finite in number, even if the domain is infinite. (This inference is quick; but as I said, this isn't a formal proof.) So we do have some justification to think that the uniform explanation thesis is true in infinite domains.

To summarize the results of this argument, if the universe is finite, then SGT is *provably* equivalent to CGT; and if the universe is infinite, then SGT is *arguably* equivalent to CGT.

7. The end?

[Again writing in the present tense.] So, now that we have the proof from the previous section, are we entitled to stop worrying about the difference between CGT and SGT? Or, to put things another way, does the contrastivity of grounding have any bearing on metasemantics?

I now think that that would be too quick. The "theorem" presented in section 6 relies on the "metaphysical assumption" (MA) that the "*real*," "*robust*," "*genuinely explanatory*" relations are closed under conjunction, and yet nonetheless fine-grained and sparse. Although not explicitly contradictory, these two ideas pull against each other.

Because these two ideas pull against each other, it is therefore quite difficult to have stable intuitions as to what the "*real*," "*robust*," "*genuinely explanatory*" relations are supposed to be. And since that is precisely what is at issue between CGT and SGT, I conclude that this issue is quite a mess.

---

[2] For another example, consider the Kripkean story as to why "Aristotle" (as used by a contemporary) refers to Aristotle rather than Plato: it is because there is a certain chain of deferential use that traces back to Aristotle rather than Plato. Well, presumably that same chain will also explain why "Aristotle" refers to Aristotle rather than Socrates, Sophocles, Empedocles, and every other ancient Greek (and person, for that matter).