Compositionality and Expressive Power: Comments on Pietroski

Elmar Unnsteinsson

Preprint. Please cite published version.
Croatian Journal of Philosophy (2020) 20(3): 295–310
Link to published version

Abstract

Paul Pietroski has developed a powerful minimalist and internalist alternative to standard compositional semantics, where meanings are identified with instructions to fetch or assemble human concepts in specific ways. In particular, there appears to be no need for Fregean Function Application, as natural language composition only involves processes of combining monadic or dyadic concepts, and Pietroski's theory can then, allegedly, avoid both singular reference and truth conditions. He also has a negative agenda, purporting to show, roughly, that the vocabulary of standard truth conditional semantics is far too powerful to plausibly describe the linguistic competence of mere human minds. In this paper, I explain some of the basics of Pietroski's compositional semantics and argue that his major objection to standard compositionality is inconclusive, because a similar argument can be mounted against his own minimalist theory. I argue that we need a clear distinction between the language of the theorist—theoretical notation—and the language whose nature we are trying to explain. The theoretical notation should in fact be as expressively powerful as possible. It does not follow that the notation cannot be used to explain mere human linguistic competence, even if human minds are limited in various ways.

*(✉) elmar.geir@gmail.com

*Many thanks to the participants and organizers of the Philosophy of Language and Linguistics Conference in Dubrovnik 2019, part of which was devoted to Pietroski's new book. Special thanks are owed to Dunja Jutronić for her help and patience. For discussions of the topics of this paper I am indebted to Paul Pietroski, John Collins, Michael Glanzberg, David Pereplyotchik, Daniel Harris, Finnur Dellsén, Una Stojnic, Ernie Lepore, Barry Smith, Anna Drożdżowicz, Michael Devitt, Paolo Santorio, and David Sosa. This work was supported by the Icelandic Research Fund (grant no. 163132).

1 Introduction

In his book, *Conjoining Meanings: Semantics without Truth Values* (2018), Paul Pietroski develops his own minimalist compositional semantics, based on only a single compositional principle which takes every complex expression to encode a monadic concept. This is hailed as a much more plausible theory of semantic competence than alternative views that trade in an infinite number of expression types and, usually, more than one principle of composition. The result is a robust and serious internalist alternative to the existing externalist orthodoxy, which Pietroski believes is too mired in the vocabulary of extensions, functions, and truth conditions.

In this paper, I offer a very rough sketch of Pietroski's positive proposal, explaining what he takes meanings to be—namely, instructions—and how they compose (Section 2). Next, in Section 3, I present his argument for thinking that standard truth conditional semantics is far too powerful to be an appropriate tool for describing semantic competence in normal human beings. I also describe one of his arguments against function application as a compositional principle, namely that it misrepresents monadic concepts as relations. According to Pietroski, the concept $F(x)$ is not a function from objects to truth values—making a monadic concept relational—but a mental device of classification, to classify things as *F*. I argue, however, that the two proposals are either incomparable, or, if they are made comparable, they may just as well turn out to be ontologically and theoretically equivalent. This all depends on further commitments, not encoded in the mere notation for function application, which remain optional for the truth conditional semanticist.

In Section 4, however, I argue that Pietroski's own semantics would be subject to objections very similar to the ones he presents against truth conditional semantics. Briefly, if we think of his own proposal in terms of basic syntactic types and some function with those types as its domain, it is easy to see how his proposal will generate a boundless number of new types. If this is right, Pietroski himself seems to assume compositional capacities that are, by his lights, too powerful to be ascribed to finite human minds. I conclude, on the contrary, that this shows that the argument itself is flawed. Roughly, the strength of the notation used by the theorist is no guide to the actual metaphysical commitments assumed in the theory itself. So, as before, I have failed to find a substantial cause for disagreement between Pietroski and his alleged opponent.

Finally, in Section 5, I describe a distinction that might be of help in this debate. This is the distinction between the representational system that the theorist uses to explain and describe some cognitive phenomenon and the system employed by the cognitive agent or agents under scrutiny. There are reasons to think that, in principle, these cannot be identical. If so, we are free to take truth conditional semantics, and Pietroski's minimalist semantics, as competing models of the same thing, rather than, say, attempted reproductions of what actually happens in human minds. And this is the natural conclusion, showing that comparing the two models may be even more difficult than Pietroski assumes.

2 Slang and Slanguages

Shaking his head over the many meanings of 'language,' Pietroski defines 'Slangs' as naturally acquirable human languages. The languages of logic, mathematics, musical notation, and so on, are not at issue. More precisely still, Slangs are constituted by a collection of mental processes that determine sets of pronunciation-meaning pairs. Presumably, sign languages are Slangs, so 'pronunciation' must also apply to the manner of producing sign expressions. Long ago, David Lewis (1975) would have defined Slangs as the sets of pronunciation-meaning pairs themselves, but Pietroski thinks this is a mistake. Still, we are all people of the same trade; semanticists working to discover the nature of the meanings or interpretations generated by the mental processes in question.

Pietroski's enduring methodological commitment is to the mantle of compositionality. The only thing we know for certain about the semantic properties of Slang expressions is that they must, must compose. Composable Meanings (CMs) are then identified as instructions for fetching and assembling human concepts. To see the point, it is best to think about what CMs cannot be. First, these meanings are not human concepts. For example, if there is a human singular concept of the person Pierre, this concept cannot be the Composable Meaning of the Slang expression 'Pierre.' Second, CMs are not extensions. For example, if the extension of the human singular concept of Pierre is Pierre himself, this entity cannot be the CM of 'Pierre.' Third, CMs are not instructions for how to use expressions or pronunciations. For example, if the pronunciation /pierre/ ought to be used to refer to Pierre, the instructions so to use it cannot be the CM of 'Pierre.'

More positively, CM-semantics involves instructions which are both composable and executionable by human minds. There are two kinds of instructions, or mental processes, *fetch*-processes and *assemble*-processes. Fetching is simple. Take the polysemous Slang expression 'book.' The Composable Meaning of 'book' is an instruction to fetch a concept stored at the 'book'-address. Since 'book' is polysemous, this particular address is home to more than one concept. Say there are two 'book'-concepts, one of books as concrete entities we can buy or burn, another of books as abstract collections of information we can choose to forget. If so, the CM of 'book' is an instruction with two admissible executions, either to fetch the concept BOOK:CONCRETE or the concept BOOK:ABSTRACT. To understand the Slang expression 'book' on a given occasion, is to execute the associated instructions to fetch either of the two concepts stored at the appropriate lexical address.

It is worth pausing to think about how meanings are being individuated on this account. Strictly speaking, 'book' has only one meaning, because meanings are instructions and the instruction for 'book' is to fetch a concept at a given address in the mental lexicon. The polysemy of 'book' consists in the fact that the address is home to two concepts and, thus, executing the instructions on a given occasion can have either of two end results. More natural, perhaps, would be to say that meanings are constituted by sets of admissible instruction-executions. That way, 'book' would have two meanings according to CM-semantics. Let's reserve the expression 'recipes' for ordered pairs of instructions and admissible executions. Polysemous expressions are expressions with more than one recipe, that is, more than one meaning. As we will see, I suspect that Pietroski must insist that Composable Meanings are instructions and not recipes. And he might very well be right.

A few words on fetching and assembling. To fetch a concept is a specific mental process which ranges over composable human concepts and is triggered by atomic expressions like 'red' and 'dot.' The lexical addresses already mentioned can only be home to composable concepts, not any old human concept. Further, these concepts can either be monadic or dyadic, but not triadic or more. Well-formed instructions to fetch a concept can themselves be composed to yield more complex instructions. Complex instructions of this sort, encoded by molecular expressions like 'red dot,' trigger the appropriate assemble-process. There are two major processes of this sort, called M-join and D-join, the results of which are unsaturated predicative concepts. So, assemble-processes take fetchable concepts and either M-join them or D-join them. 'Red dot' is M-joined to yield the concept RED DOT(_), applying to red dots. An expression like 'above' fetches a dyadic concept, ABOVE(_, _) with two unsaturated argument places. This concept may be composed with RED DOT(_) with the D-join operation, to yield a monadic concept applying to whatever is above a red dot, roughly ABOVE RED DOT(_). Two dyadic concepts cannot be joined, only two monadic ones or a monadic and a dyadic one. And the result is always a composable monadic concept. Here I suppress a number of important details, for example about how exactly D-join works, because they are not important for the points I wish to make.

I should stress that, for CM-semantics, it is entirely possible that humans possess and regularly employ unfetchable and non-assembling natural concepts. The Composable Meaning of 'Pierre' is an instruction to fetch a composable concept at the 'Pierre'-address. According to Pietroski, we better think of this as a special monadic concept, PIERRE(_), applying to objects which are called 'Pierre' (p. 249). This is what makes the concept composable via the assemble processes of M-join and D-join. Even so, the CM-semanticist can very well allow that hearing the expression 'Pierre' uttered on a given occasion may come to activate another concept in one's mind, namely the singular concept PIERRE (p. 108). But activation and fetching are not the same thing. Meanings may activate a number of mental phenomena, some even regularly and reliably, but, still, these mental phenomena will not be composable concepts.

Or, so the story goes at least. The Fodorians in the room might well wonder what, if anything, can be meant by the idea of a concept that does not compose. If there is a language of thought ('Mentalese'), presumably it enjoys a compositional semantics. So, if there is a singular concept of Pierre, the denoter-concept PIERRE as it were, then it must compose with other concepts, like the concept SNORES , to yield a proposition with a truth condition, being true if and only if Pierre snores. If we think in a compositional Mentalese language—which, we should remember, need not be a Slang—the theory of CM-semantics may appear inherently unstable. Put it this way, assuming that Mentalese contains the purely denotational expression 'Pierre' whose content is exhausted by the individual, Pierre, and thus, that such Mentalese expressions must compose (for Mentalese is compositional), why shouldn't Slangs be able to contain such expressions too? If Mentalese-'Pierre' can fetch singular concepts—because they are compositional—why can Slang-'Pierre' not do so as well?

Not to dwell too long on this point, I believe Pietroski would be best served either to deny that there are any singular terms in the standard Kripkean sense, or to deny that we think in a compositional language of thought. As far as I can see, both may be sources of prevarication on his part. Understandably, as both are non-negotiable for some theorists.

This is my rough sketch of Pietroski's positive proposal, leaving out a lot of fascinating detail, but he also has a more negative agenda. He argues that this minimalist theory is in tension with standard possible worlds truth conditional semantics and, further, that the latter was a bold conjecture best consigned to the flames. In the next section I will focus on one particular argument in this vein, suggesting that it is inconclusive as it stands. First, however, I think it may be helpful to state more flat-footedly what seems to be the real difference between maximalist semantics—as we might call the alternative—and Pietroski's minimalism.

The minimalist rejects function application as a basic compositional principle, putting in its place something more like predicate modification (assemble processes). Both of these principles are tentatively endorsed in the standard textbook of maximalist semantics (Heim & Kratzer 1998), where the elimination of predicate modification is perceived as desirable but ultimately undoable. Heim and Kratzer do not seriously consider the option of eliminating function application. Pietroski's contribution, among many other insights, is at least that of presenting a robust theory on which only something like predicate modification is assumed. Ultimately, however, I think one question is left hanging in the air: Has it been shown, or even made plausible, that there would be a distinction at the level of mental mechanisms between a system with maximalist psychology and a system with a minimalist one? That is, is there any cognitively realized distinction between systems characterized by the assemble function and those characterized by function application together with predicate modification? As far as I can see, but certainly Pietroski disagrees, these may well turn out to be equivalent models of the same cognitive phenomenon, depending on the precise commitments of the maximalist theory. If so, then, as I will argue in the next section, we should stick to the more powerful theoretical vocabulary of maximalist semantics.

Finally, more sweepingly, minimalists need stronger arguments to eliminate extensions, truth conditions, truth values, and possible worlds from any semantic theory, even their own. Pietroski does not rule out that Mentalese has a maximalist semantics (pp. 84–85). But he needs to show that this is not the case for, otherwise, the argument for having one type of semantics for both Slangs and Mentalese come knocking. That's not the only problem, however, because if Slangs are used to activate Mentalese to, in turn, activate propositional attitudes, the connection between the two seems far too tight to argue that Slang-'Pierre' does not function, somehow or other, to fetch or activate whatever Mentalese-'Pierre' is supposed to fetch or activate. Minimalism can carve out semantic space which excludes extension and truth conditions, but the carving itself does not show that such phenomena do not exist or that they are irrelevant to the study of meaning and content. To carve things up in this way is merely to insist that meanings are instructions and not recipes, as these terms were defined above.

2.1 The Fregean Hierarchy and Classification

Pietroski indeed presents one very ambitious argument of this sort. If sound, it seems, it would establish precisely that neither Mentalese nor Slangs could possibly enjoy a maximalist semantics. The conclusion of the argument is, roughly, that the notational machinery—specifically the Fregean Hierarchy of Types and the lambda-calculus—in which maximalism is entwined, is far too powerful and sophisticated to represent the mental processes of finite human minds. In what follows, I focus on this argument, which, although impressive and intriguing, is inconclusive.

Maximalist semantics is standardly introduced by defining so-called Fregean or Montagovian semantic types (e.g., Dowty et al. 1981 , Heim & Kratzer 1998) and Pietroski points out that the types are 'boundlessly many' (p. 127). The types are defined recursively as follows:

(1) e is a type

(2) t is a type

(3) if a and b are types then $\langle a, b\rangle$ is a type

This definition represents an infinite class of syntactic or semantic categories. It follows, for example, that $\langle e, t\rangle$, $\langle\langle e, t\rangle, t\rangle$, and $\langle\langle t, t\rangle, \langle t, t\rangle\rangle$ are all Fregean types. Expressions of type e refer to objects like Pierre or Luang Prabang. Expressions of type t stand for truth-values, True or False. Slang expressions like 'snores' are recursively defined functions from objects to truth-values. So, 'snores' is a type of expression which takes (the semantic value of) an e-type expression and outputs a truth-value. From this we get the intuitively correct result that 'Pierre snores' is true if and only if the function encoded by 'x snores' outputs truth when its argument place is assigned the object referred to by 'Pierre.'

Pietroski points out that we can define levels of complexity for the Hierarchy of Types. (1) and (2) in the recursive definition is Level Zero. Level One is defined as the set of possible functions $\langle a, b\rangle$ whose members a and b are at Level Zero. So, Level One is $\langle e, e\rangle$, $\langle e, t\rangle$, $\langle t, e\rangle$, and $\langle t, t\rangle$. Each higher level is defined in the same way, such that Level N builds types from all levels lower than N. Level Zero has two types, One has four, and Two has thirty-two (for example, $\langle\langle e, t\rangle, t\rangle$). When we reach Level Four, we have more than two million types. Many types at Levels Three and Four seem bizarre and are not realized, Pietroski states, by human minds. But we certainly can define such types, and some might be important in logic or mathematics, like the Fregean concept of an ancestral, which is Level Four (p. 128). Pietroski wants to conclude that Slangs cannot be Fregean languages, because they are far too powerful to be realized in the mind of every human who speaks a natural language. If this follows for Slangs, it follows for Mentalese as well. Fregean thoughts are for Frege, not us mere mortals.

Maximalists semanticists also use the extremely powerful lambda-calculus to make good on their promise of a truth conditional semantics. As Pietroski shows, the Fregean idea, which required Tarski's notion of quantification over variant sequences of objects, is to treat monadicity as a special case of relationality (p. 83). For example, the semantic value of 'brown' is a special kind of truth function defined on the basis of lambda-abstraction, $\lambda x \langle x \text{ is brown} \rangle$, mapping x to Truth if x is brown and to Falsity otherwise. So, this is a function, roughly, taking one set of objects to T and another set to F, determining sets of things satisfying a certain condition, namely having the property of being brown. In principle, this allows maximalists to model any Slang expression which is not of type *e* or *t* as a mathematical function. This is implicit in the Fregean Hierarchy already discussed, where 'brown,' for example, is an expression of type $\langle e, t \rangle$. But the lambda-calculus is necessary to make this idea coherent and workable.

To the contrary, Pietroski argues, monadic concepts like BROWN(_) do not represent functions at all, especially not a sophisticated lambda-function from objects to truth-values. Rather, he claims, such concepts are classificatory; they let us classify objects without relating them to truth values (p. 83). He explains that there is a psychological distinction between relational and classificatory concepts, the former are for '... classif[ying] things, into those that meet a certain condition (e.g., being a rabbit) and those that do not. Anything that meets the condition satisfies the predicate, which applies to anything that meets the condition.' (p. 28). Relational concepts, however, are ones like ABOVE(_, _), where one object is related to another and the relation is satisfied when the objects are in fact, in this example, related such that one is above the other. Pietroski's point is that BROWN(_) is simply not relational and thus it is misrepresented by the Fregean theory.

But what is the difference exactly? If we tried to construct a creature whose monadic concepts are non-relational because they are merely classificatory, what would we get? It seems almost impossible to end up with anything other than a relational concept in Frege's original sense. This is because classifying things into those with the property *F* and those that do not have the property *F* is equivalent to classifying things into those that are truly judged to be *F* and those that are falsely judged to be *F*. At least, that seems to be Frege's position, as Sanford Shieh

(2019) has recently tried to show in some detail. Briefly, Shieh's interpretation is that Frege believed that the truth of the thought that *p* is constituted by the obtaining of what the thought represents. To recognize that the referent of '*a*' falls under the referent of 'is *F*' is thereby to recognize that the thought that a is *F* refers to the True (2019, 108). If this is right, there is no difference at all between classificatory concepts and Fregean concepts with a single unsaturated argument place, indicated by '*x* is *F*' (Glanzberg 2014, 267 makes a related point I think).

Still, I don't think this objection is conclusive. What we would need is some account of the difference between classifying and relating to truth values and, in fact, I think there might be a plausible account of this sort. But it is not one I can find in Pietroski and of course that might be my own fault. Anyway, very roughly, if we hold that non-declarative clauses, like imperatives and interrogatives, do not relate objects to truth values, because they have no truth conditions, then the properties or concepts occurring in those clauses must classify objects without relating the objects to truth values. But the very same properties can occur in declarative clauses too, and, so, we have something of puzzle. As I understand most proposals in the literature on imperative semantics, for example, they ultimately model properties—even as they occur in imperatives—as functions from objects to truth values (see, e.g., Roberts 2018). I hope to address this issue in future work, so I leave it unresolved here.

As if by the law of gravity, this discussion is veering dangerously close to the whole issue of the connection between logic and psychology, and Frege's own complicated view of that connection. The topic is too massive, controversial, so we will mostly steer clear. But Pietroski writes that the relational conception of monadic concepts, and the Function-Argument structure of thoughts more generally,

> …led to a brilliant conception of how thoughts could be logically related. But like Frege, I don't think it should be viewed as a psychological hypothesis, even if it can be viewed as a model for a certain kind of ideal cognition. (p. 83)

It would seem, however, that Frege's anti-psychologism is so staunch as to exclude any psychology whatsoever, no matter how powerful. His anti-psychologism is driven by the conviction that logic is an autonomous science, not reducible to (anyone's) psychology. Thoughts are timelessly and mind-independently true or false. Ditto, then, for the logical relationships between different thoughts (Frege 1918/1956).

Still, undeniably, thoughts in Frege's sense are relevant to psychology, because they are what we grasp, judge, and assert. Even more, they are what any cognitive creature, ideal or not ideal, would grasp, judge, or assert, in trying to discover true thoughts, which for Frege is the 'work of science' (Frege 1918/1956). If ideal cognition discovers the true thought that p, and the truth of p consist in the fact that a is F, then no one else can discover exactly the same thing without recognizing that a is F. For Frege, a model of ideal cognition is a model of cognition. So, if the Function-Argument structure reveals to us (the theorists) how certain thoughts are related logically, we have also found out which particular thoughts need to be grasped by anyone credited with recognizing those relations. The structure itself would then seem to be theoretically indispensable.

Finally, the quote may help us better to diagnose the problem of distinguishing classificatory and relational concepts. As Pietroski says, there is a sense in which Frege does not intend the relational analysis as a psychological hypothesis. What this means for Frege, I take it, is that the structure imputed on the thought by the logician has no direct psychological relevance. That is, it does not tell us anything about the nature of the mental acts of grasping, judging, or asserting; except insofar as those acts are individuated in terms of their objects. More importantly, though, if the Fregean structure is not a psychological hypothesis in this sense, then there is no tension between that structure and any other structure, when it comes to human psychology. The relational conception of monadic concepts, and monadic thoughts like a is F, can very well model whatever it is that the classificatory conception models. At the level of psychological mechanism, the two are equivalent. If both range over exactly the same set of thoughts—if we are allowed the Fregean notion of a thought—they are completely equivalent.

3 Notation and Expressive Power

Pietroski subscribes to the following methodological strategy:

> In defending any proposal about meanings, one must also take care to not assume implausibly powerful expression-generating capacities. Similarly, in defending any proposal about syntactic structures, one must take care to not assume implausibly powerful operations of semantic composition. (p. 294)

Surely, the counsel is sound. But how exactly do we judge degrees of power?

And when we know how, will it really follow that Pietroski's minimalism has the right degree and maximalism not? My response to the first question is, one, that at least it is not judged by the expressive power of the notation employed by the theorist and, two, that any theory powerful enough to describe the basic range of human semantic competence, has exactly the right degree of expressive power. It may, still, lack explanatory power, which is a different matter. My response to the second question is that, no, this particular line of argument against maximalism is inconclusive, at least if both parties are working with some distinction between competence and performance.

There appear to be two notions of power in play here, productivity and composability. A theory's degree of productivity is determined by the number of new items, or syntactic types, it can generate from some finitely stateable base of principles. Generative theories of semantic competence are normally thought of as generating a potential infinity of new items. The degree of composability, however, is determined by the number of new items or types which can serve as inputs into the compositionality function postulated by the theory. Presumably, if the items are infinite in number, the domain of the compositionality function will be infinite as well.

The primary virtue of Pietroski's compositional principles is their simplicity. The procedures modeled by these functions are dumb and can certainly be performed by specialized cognitive mechanisms, rather than needing the resources of a full-blown mind. But this virtue is shared by function application.

To see this more clearly, consider the productivity and composability implicit in the notational systems we unhesitatingly, and rightly, employ in describing the thoughts of young children. We posit that young children can represent the thought that John thinks that cows are brown and that 2 + 2 = 4. In the first case, we employ the incredibly powerful formula '*A* thinks that *p*.' By simple recursion, we will get '*A* thinks that *B* thinks that *p*.' So, postulating thoughts about others' propositional attitudes involves notation which is infinitely productive and boundlessly compositional. Similarly, in the second case, we employ the concept of addition, symbolized by '+'. In particular, we attribute the thought that 2 + 2 = 4 to children well before they develop the capacities to represent numbers higher than 100. If Pietroski is right, this should be a problem. The symbol '+' is infinitely productive and boundlessly compositional and thus should not be used to describe the thoughts of young children. But that would be the wrong conclusion to draw.

Moreover, CM-semantics can be developed recursively as a Hierarchy of Types, just like the maximalist theory (Pietroski mentions this possibility briefly on p.

113). CM-semantics has two principles of composition, M-join and D-join, both of which will have monadic composable concepts as their outputs. In M-junction, if $\Phi(\_)$ and $\Psi(\_)$ are composable concepts then $\Phi(\_)\hat{}\Psi(\_)$ is also composable. Here '^' stands for the compositional process of conjunction, let's call it 'junction.' In D-junction, if $\Phi(\_)$ and $\Psi(\_,\_)$ are composable concepts then $\Psi(\_,\_)\hat{}\Phi(\_)$ is a composable monadic concept, for example the concept for classifying things as being above a red dot. A brown cow above a red dot would be an example where the results of M-junction and D-junction are themselves joined by M-junction. Let's now define the Pietroskian Hierarchy by recursion:

(1) *m* is a type

(2) *d* is a type

(3) if *a* and *b* are types then $(a\hat{}b)$ is a type

Here, *m* stands for monadic concepts and *d* for dyadic concepts. The peculiarities of junction ('^') are important, as it behaves very differently from the set theoretic notion of an ordered pair employed by the maximalist. With that in mind, if Level Zero of the Hierarchy is exhausted by m and d, Level One is exhausted by the two types $(m\hat{}m)$ and $(m\hat{}d)$. Level Two is only slightly more complex, if defined in terms of the number of added admissible junctions we can perform by the junction of types at Levels Zero and One, e.g., $[m\hat{}(m\hat{}m)]$ and $[(m\hat{}d)\hat{}(m\hat{}d)]$. As soon as we reach Level Three we have 56 types, Level Four has 2,212, and Level Five has 2,595,782 types. We have clearly outstripped the finite cognitive capacities of mere human minds. Still, Pietroski is happy to employ notation which belongs to Level Four (p. 202), if interpreted as $[(d\hat{}m)\hat{}m]\hat{}[m\hat{}(d\hat{}m)]\hat{}m$:

$$\exists[\textsc{agent}(\_,\_)\hat{}\textsc{al}(\_)]\hat{}\textsc{chase}(\_)\hat{}\textsc{past}(\_)\hat{}\exists[\textsc{patient}(\_,\_)\hat{}\textsc{theo}(\_)]\hat{}\textsc{gleeful}(\_)$$

This would be a proposed logical form for 'Al chased Theo gleefully,' making explicit who is the chaser (agent) and who the chasee (patient). I have not explained Pietroski's notion of existential closure, '∃,' but it is needed to coordinate one of the gaps in the *d*-type with the gap in the *m*-type. So, roughly, *x* is above a red dot if there is a red dot *y* such that *x* is above *y*. Notice, also, that in CM-semantics, juncture-processes are insensitive to order, so $[(d\hat{}m)\hat{}m]$ is equivalent to $[m\hat{}(m\hat{}d)]$.

Why does Pietroski not see this as a problem for his minimalism? That is, why does he assume that this form of argument is a blow to maximalism and not to his own theory? I think there are two reasons which, although illuminating, are inconclusive.

First, in considering something like this point, Pietroski seems to assume that (3) is not really true. That is to say, he believes CM-semantics has only two syntactic types, *m* and *d*, and claims that all complex expressions would be of type *m*. Therefore, there are no non-basic types in his semantics (p. 113). I do not see how this could be true. Junction is a process or operation modelled by a function. The value or output of the function is always an expression or concept of type *m*, but that doesn't mean that junction itself is of type *m*. That would be like saying that $\langle e, t\rangle$-expressions are really *t*-type expressions, because their output is always a truth-value. Rather, junctions are functional expressions, whose domain could be defined as follows. Take the set *D* consisting of (i) all items of type *m*, (ii) all sets *m*, *d* of items of type *m* and *d*, and (iii) the infinite number of sets resulting from every possible combination and repetition of the elements in (i) and (ii). Now remove every non-set from *D*, that is, delete the element *m*. Next, take the set *G* of all items of type *d* and define the Cartesian product of sets *D* and *G*, $D \times G$ (so we get, for example, $\{\{m, d\}, d\}$). Call the resulting set *F*. At last, the junction-function is defined as a mapping from every element in *F* to a set of expressions or concepts of type *m*. For example, this function takes the pair BROWN(_) and COW(_) as input and yields the monadic concept BROWN(_)^COW(_), which classifies things as brown cows. *F* contains infinite sets, so it goes without saying that human minds are unable to apply junction to every member of *F*.

Similarly, maximalist semantics really contains three types, *e*, *t*, and *f*, where *f* is a function defined in terms of the first two types. CM-semantics contains expressions of type *m*, *d*, and *j*, where *j* is a function from *F* to a set of *m*-types. This way of looking at the matter is unavoidable, because otherwise we will lack the resources to distinguish the composed concepts $\Phi(\_)\hat{}\Psi(\_)$ from the non-composed composable concepts $\Phi(\_)$. Both are indeed composable, but only the former is composed with the operation of junction. We can think of *j* as a function from, for example, $\{\{m, d\}, m\}$ to $[(m\hat{}d)\hat{}m]$, where the latter is itself a monadic concept of a certain sort. If this is right, junction is both infinitely productive and infinitely compositional. To be clear, I do not hold that this is a problem for CM-semantics, to the contrary. My point is only to argue that if too much notational power is a problem for maximalism, for the reasons Pietroski adduces, it is also a problem for minimalism. The relative power of the notation is irrelevant to the basic metaphysical questions at issue.

Second, Pietroski seems to shy away from relying explicitly on a distinction between competence and performance. But, certainly, some such distinction must be working in the background. This could easily be applied to explain how

CM-semantics can appeal to potentially infinite operations like juncture without ascribing super-human cognitive powers to humans. Semantic competence, as far as composable meanings go, is properly described in potentially infinite terms, if the base is finitely stateable. If this strategy is available to the minimalist, however, it must also suggest itself to the maximalist. This would have stopped the objection from too much notational power in its tracks. Maximalists are free to argue that the only types ever required in actual linguistic performance will remain below Level Four. Furthermore, large swaths of the types represented at the Levels above Zero will never be used at all. Level One, for example, would have it that there are expressions of type $\langle e, e \rangle$, which are not needed for the semantics of Slangs.

When both theories—minimalism and maximalism—have access to a distinction between competence and performance, direct comparisons of degrees of complexity become more cumbersome. Possibly, still, Pietroski's formulation of different Levels in the Hierarchy of Types could remain useful. But the comparison would have to be between different analyses of one and the same sentence, spelling out the varying operational power required by each analysis. However, if function application and classification are not really different operations, which is still undecided I think, it is not clear which theory wins out on simple sentences. 'Joe snores' is an instruction to fetch two monadic concepts and to M-join them. The resulting monadic concept classifies things into the set of Joe-snoring elements and the set of non-Joe-snoring elements (objects, worlds, or situations). For the maximalist, perhaps, it is an instruction to mentally fetch Joe himself and apply the snoring-function to him, again dividing things into the Joe-snoring elements and everything else. I must admit, in light of the number of different theoretical commitments still open to both theorists, that I have a hard time judging which theory involves more powerful mental operations. All I can say is that focusing on individual sentence or clause types may make the project tractable.

4 The Language of the Theorist

To recap, it seems more difficult to eliminate function application than Pietroski assumes, for the notion is, at least potentially, equivalent to his own notion of classification. We would need a story on which function application and classification must have different realizations in mental mechanisms. Moreover, compositional semantics will always call for notation which is in principle boundlessly powerful, and a competence-performance distinction to explain why only a small part of those representational resources are needed to describe the facts of human performance. Even CM-semantics traffics in notation which allows for boundless productivity but, still, the theory is not exclusively concerned with ideal minds.

It is reasonable to conclude that the most powerful notation on offer is the best one to go by, as long as we can find no hypotheses about cognitive implementations hidden in the notation itself. If so, the Fregean Hierarchy of Types and the lambda-calculus are very good bets. Even if these are powerful, they can be used to describe simple, dumb operations, like applying *Fx* to some object *a*, in such a manner that the proposition entertained is true if and only if *Fa*.

As I have tried to argue elsewhere (see my 2016, 2019, and forthcoming), we should make a distinction between the representational system employed by the theorist to describe and explain the cognition and actions of particular human agents, and the representational system employed by those agents themselves. Human thinkers are often confused in ways that make it inadvisable to simply incorporate their own representations into our explanations. Of course, these representations can always be mentioned, quoted, or otherwise referred to, but they cannot be used directly by the theorist. A simple example is a thinker who confuses the identical twins Bill and Biff and has only a single concept, labeled 'John,' to represent both. Strictly speaking, the thinker then lacks the representational resources to think explicitly about Bill without thinking about Biff at the same time. But still, I have argued, we must be able to ascribe the false belief that Bill is identical to Biff to this agent. This must then be some form of implicit belief.

An opposing view would be that, as theorists, we must be able to reproduce or mirror the confused mental state in our theoretical vocabulary. Well, yes, we must be able to refer to those mental states somehow, but we must also be able to ascribe beliefs which the agent is constitutively unable to represent explicitly. Stuart Hampshire (1975, 123) articulated these two options clearly: 'Perhaps the confusion in his mind cannot be conveyed by any simple account of what he believes: perhaps only a reproduction of the complexity and confusion will be accurate.' So, either our descriptions are reproductions of the blooming, buzzing confusion of our inner lives or they are, rather, models. But if we must sometimes posit the belief that p for purposes of explanation even when the explicit thought that *p* is unavailable, we have in effect given up on reproductions.

CM-semantics is in business of providing a model of human minds, just as much as the maximalist alternative. Maximalists can coherently accept any performance-restriction that the minimalist cares to propose. For example, they could agree that human minds can only compose two unary relations or one unary and one binary relation. Nothing in the notation itself disallows the restriction. The real nub of the argument is whether we can find deep, metaphysical differences between conjunctive or predicative composition and function application. Even if I lean one way on this issue right now, I genuinely think it is an open and interesting question.

5 Conclusion

One of my underlying themes has been to suggest that, sometimes at least, objections against a compositional semantics for Slangs should also be objections to the same compositional semantics for Mentalese. So, if the objection would have absurd or unpalatable consequences for Mentalese, perhaps the objection is not reliable in general. Thus, if the objection can be resisted for one of these it can also be resisted for the other. It bears emphasizing, however, that this point certainly depends on various assumptions, one of which is the very notion that we think in a compositional Mentalese.

At one point, Pietroski argues that the Liar paradox is a problem for truth conditional semantics (Chapter 4). Roughly, if my favorite sentence can be 'My favorite sentence is not true,' standard truth conditional semantics will involve contradictions. Pietroski wants to infer that truth conditional semantics should go. If this is correct, it would suffice to prove that Mentalese cannot have a truth conditional semantics either, because my favorite thought can be that my favorite thought is not true. But it just seems too incredible to believe that our thoughts cannot be true or false, and thus have truth conditions. Perhaps we should conclude, then, that there is something wrong with thoughts or sentences of this kind, not that neither thoughts nor sentences can be true or false.

The broader point is to suggest that the expressive power of our theoretical vocabulary is not, as such, a reliable indicator of the explanatory power, or lack thereof, of any theory expressed by that vocabulary. Working formal semanticists tend to avoid making proclamations about where human linguistic competence ends and god-like mental powers would have to start. But surely, Pietroski is correct to point out that Level Four in the Fregean Hierarchy of Types is not needed to explain competence in Slangs. My argument is, basically, that one need not be a CM-semanticist, that is, one need not eliminate function application, to make this particular point.

Bibliography

Dowty, D. R., Wall, R. E., & Peters, S., 1981. *Introduction to Montague Semantics.* Springer.

Frege, G., 1918. "Der Gedanke. Eine logische Untersuchung." *Beiträge zur Philosophie des deutschen Idealismus*, 1:58–77.

Glanzberg, M., 2014. "Explanation and Partiality in Semantic Theory." A. Burgess & B. Sherman (eds.), *Metasemantics: New essays on the foundations of meaning*, Oxford University Press, pp. 259–292.

Hampshire, S., 1975. *Freedom of the individual.* Chatto & Windus, 2 edn.

Heim, I. & Kratzer, A., 1998. *Semantics in generative grammar.* Blackwell.

Lewis, D., 1975. "Languages and Language." K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science*, University of Minnesota Press, pp. 3–35.

Pietroski, P. M., 2018. *Conjoining Meanings: Semantics Without Truth Values.* Oxford University Press.

Roberts, C., 2018. "Speech Acts in Discourse Context." D. Fogal, D. W. Harris, & M. Moss (eds.), *New Work on Speech Acts*, Oxford University Press, pp. 317–359.

Shieh, S., 2019. *Necessity Lost: Modality and Logic in Early Analytic Philosophy, Volume 1.* Oxford University Press.

Unnsteinsson, E., 2016. "Confusion is Corruptive Belief in False Identity." *Canadian Journal of Philosophy*, 46(2):204–227.

———, 2019. "The Edenic Theory of Reference." *Inquiry: An Interdisciplinary Journal of Philosophy*, 62(3):276–308.

———, forthcoming. *Talking About.*