Essential Meanings

Abstract: An entity is *semantically neutral* just in case it lacks inherent meaning, content, or semantic properties. An entity is *semantically permeated* just in case it has its meaning essentially. The distinction between semantic neutrality and permeation applies to natural language, mental representation, and formal logic. As emphasized by philosophers such as Plato, Ockham, Frege, Putnam, Fodor, and Burge, fundamental questions turn on whether the relevant domain features semantically neutral or semantically permeated entities. I try to clarify this crucial distinction by offering a modal analysis. Deploying my modal analysis, I elucidate how the semantically permeated approach to mental representation accommodates *syntactic structure*.

§1. Semantic permeation and semantic neutrality

A traditional view, propounded by Aristotle and extensively developed by Ockham, holds that natural language words are inherently meaningless, while concepts or mental representations have essential semantic properties. As Ockham puts it,

> [a] concept or mental impression signifies naturally whatever it does signify; a spoken or written term, on the other hand, does not signify anything except by free convention. From this follows another difference. We can change the designation of the spoken or written term at will, but the designation of the conceptual term is not to be changed at anybody's will.[1]

This traditional view has two core elements. First, the connection between a word and its meaning is arbitrary, barring a few possible exceptions such as onomatopoeia. In general, a word

---

[1] *Summa Logicae*, I.1, in his *Philosophical Writings, A Selection*, P. Boehner (ed and trans), (London: Nelson, 1957).

has no inherent bias towards one meaning over another. Second, the connection between a concept (or mental representation) and its meaning is not arbitrary. Rather, it is essential to the concept (or mental representation) itself. As Burge puts it, "a concept of an eclipse could not be the concept that it is if it did not represent, or it were not about, eclipses."[2]

I will say that an entity is *semantically neutral* just in case it bears an arbitrary relation to its meaning (assuming it even has a meaning). A semantically neutral entity lacks inherent meaning, content, or semantic properties. It could have had a different meaning, or no meaning at all, without any change in its fundamental nature, identity, or essence. I will say that an entity is *semantically permeated* just in case it has its meaning essentially. We cannot change the entity's meaning while holding fixed its fundamental identity, nature, or essence. In my terminology, the traditional view holds that natural language words are semantically neutral but that concepts are semantically permeated. This view continues to find advocates, such as Burge.

Critics have attacked both components of the traditional view. Regarding the first component, a few theorists argue that proper analysis of natural language should recognize semantically permeated words instead of, or in addition to, semantically neutral ones. More specifically, some philosophers beginning with Plato's Cratylus insist that words have their referential properties essentially.[3] The second component of the traditional view, that concepts or mental representations are semantically permeated, encounters sustained opposition within contemporary philosophy of mind. An extreme example is Stich, whose *syntactic theory of mind* exhorts cognitive science to jettison semantics altogether.[4] On Stich's view, scientific psychology should ignore content, treating mental processes as computations defined over mental items described in purely syntactic, non-semantic terms. Many philosophers who reject

[2] T. Burge, *Foundations of Mind* (Oxford UP, 2007), p. 292.
[3] A recent example: J. Justice, "On Sense and Reflexivity," *The Journal of Philosophy* 98 (2001), pp. 351-364.
[4] S. Stich, *From Folk Psychology to Cognitive Science* (MIT Press, 1983).

Stich's extreme disavowal of semantics share his predilection for semantically neutral mental items. For instance, Fodor argues that propositional attitudes are relations to mental representations drawn from a semantically neutral *language of thought*, or *Mentalese*.[5] Fodor holds that Mentalese words have semantic properties, but he denies that they have such properties essentially. A Mentalese word might have had different semantic properties, or no semantic properties at all.

The distinction between semantic neutrality and semantic permeation also arises within logic. Should artificial languages constructed by logicians feature items that lack any inherent meaning, or should they feature inherently contentful items? As van Heijenoort observes, proponents of "logic as calculus" (*calculus ratiocinator*) favor the first approach, while proponents of "logic as language" (*lingua characteristica*) favor the second.[6] The issue figures prominently in the Frege-Hilbert correspondence, with Frege embracing semantically permeated artificial languages and Hilbert favoring semantic neutrality.

Clearly, the distinction between semantic neutrality and semantic permeation is central to the study of language, mind, and logic. It is also rather murky. What does it mean to say that a representation has, or does not have, its meaning essentially? What is at issue when we ask whether words/concepts/mental representations have inherent semantic properties? What are logicians debating when they debate whether to employ inherently meaningful formal languages? Philosophers have devoted surprisingly little attention to these questions. I will try to rectify the omission by offering a rigorous analysis of "semantic permeation" and "semantic neutrality." I will then deploy my analysis to explore the prospects for semantically permeated approaches to mental representation.

---

[5] J. Fodor, *LOT2* (Oxford UP, 2008).
[6] J. van Heijenoort, "Logic as Language and Logic as Calculus," *Synthese* (17) 1967, pp. 324-30.

§2. A modal analysis

The literature features diverse conceptions of meaning or content: Fregean senses, Russellian propositions, sets of possible worlds, and so on. I will handle this diversity by speaking about "semantic values," making no assumptions regarding what those are. I write $val(x, d)$ to mean that d is a semantic value of x. I do not assume that x has a unique semantic value, although for many purposes that assumption might be fitting. My approach is still not ideally general, because some, such as Davidson, decline to associate predicates and logical connectives with semantic values.[7]

Setting aside what "meanings" are, what is it for an item to have its meaning "essentially"? The orthodox Kripkean analysis of essence holds that property F is essential to t iff F holds of t in all those possible worlds in which t exists, i.e.

(1) $\Box(\exists y(y = t) \rightarrow Ft)$

In this spirit, we might say that t is semantically permeated iff

(2) $\exists d \Box(\exists y(y = t) \rightarrow val(t, d))$

leaving open whether the relevant modality is metaphysical, mathematical, or logical. However, I will now argue that (2) is not quite adequate as it stands.

Linguistic expressions, concepts, and mental representations are *types*. The primary reason we cite them is to *taxonomize* or *type-identify* other entities, where "entity" is construed broadly to include states, processes, and events. The entities so taxonomized are *tokens*. Possible tokens include inscriptions, patterns of pixel activation, acoustic events, bumps of Braille, electrical transmissions of Morse code, patterns of neural stimulation, and so on. We posit types so as to classify tokens. Metaphorically: a type's whole being is intertwined with its role in type-

[7] D. Davidson. *Inquiries into Truth and Interpretation* (Cambridge UP, 2001).

identification. (2) does not mention that the relevant entities *are* types, let alone illuminate the relevant type-token relation. Through this omission, (2) overlooks what is most crucial about semantic permeation: namely, that a type's tokens must share a uniform semantic value. (2) is consistent with claiming that t's tokens lack any semantic properties, or even with claiming that t is not the sort of entity that has tokens.

These reflections suggest that we replace (2) with the following definition: type t is semantically permeated iff

(3) $\exists d \, \Box \, \forall x(x \text{ is a token of } t \rightarrow val(x, d))$

(3) does not entail that a token of type t has a uniform semantic value in all worlds, or even in all worlds where the token exists. In other words, (3) does not entail

(4) $\forall x[x \text{ is a token of } t \rightarrow \exists d \, \Box \, (\exists u(u = x) \rightarrow val(x, d))]$

The type has its semantic value essentially, but individual tokens need not. A token of t might arguably have had wildly different causal, functional, or teleological properties, in which case it would doubtless have had a different semantic value. (3) demands only that, in such a case, the token would not have had type t.

I have ignored a wrinkle: temporal change. Plausibly, word-types (e.g. "gay") can change their meaning over time. It is less clear that word-*tokens* can change their meanings over time. If they can, we might include a temporal parameter in *val*, which would complicate our definitions. Alternatively, we might treat "$\Box$" as implicitly quantifying over temporal indices.

Another wrinkle concerns *non-actual* semantic values. One might hold the following combination of views: linguistic expressions are abstract entities, so they exist in all possible worlds; but some linguistic expressions refer to objects (e.g. Pegasus) that do not exist in the actual world. Of course, one might also contest this combination of views. But our definitions should not take a stand. Thus, it seems appropriate to replace (3) with

(5) $\Diamond \exists d \Box \forall x(x \text{ is a token of } t \rightarrow val(x, d))$

(5) only captures our intuitive starting point if we assume that the accessibility relation on possible worlds is symmetric. Only then does (5) entail that all actually existing tokens of *t* actually have semantic value *d*. Thus, (5) is a suitable definition only if we assume the *Brouwerische* axiom $p \rightarrow \Box\Diamond p$, which follows readily from symmetry.

(5) is my official definition of semantic permeation. To simplify matters, however, I henceforth assume that our quantified modal logic involves a fixed domain of objects, rather than varying domains associated with each possible world. Thus, I assume that all objects relevant to my discussion exist in all possible worlds. Under this assumption, we may replace with (5) with (3), and we may eschew the *Brouwerische* axiom. Relaxing my existence assumptions would not affect my main points, but it would clutter the exposition.

We must distinguish (5) from the familiar notion of *rigid designator*, which also relates content and modality. A *rigid designator* designates the same object in all possible worlds where the object exists, and it does not designate another object in any other possible world. When evaluating if a term is rigid, we hold its meaning fixed. We do not consider what meaning the term has in various counterfactual circumstances and then ask what its denotation would have been had it had that meaning. Instead, we ask what the term, *as used by us with its current meaning*, denotes in various counterfactual circumstances. In contrast, (5) considers the semantic values of a term's tokens *as used in various counterfactual circumstances*.

To illustrate, assume with Kripke that proper names are rigid designators. If *s* is a linguistic type that serves as a proper name in our linguistic practice, then *s* is a rigid designator *in any counterfactual situation where it has its actual meaning*. However, this is perfectly consistent with there being some counterfactual circumstance in which *s* is used with a completely different meaning, perhaps not even as a singular term. For instance, if we individuate *s* solely by its phonological properties, then tokens of *s* might have any semantic value we desire. Thus, the fact that a singular term *as used by us* is rigid does not entail that the term is semantically permeated.

Conversely, a semantically permeated singular term need not be rigid. Suppose we adopt Lewis's framework, on which an *intension* for a singular term is a function from indices (which may include a possible worlds parameter, a contextual parameter, etc.) to things.[8] For any intension ϕ, we can posit a singular term *s* that expresses ϕ by its essential nature:

$$\Box\ \forall x(x \text{ is a token of } s \rightarrow val(x, \phi))$$

Then *s* is semantically permeated, but it may not be rigid. It is rigid only if ϕ's output does not vary with an index's possible worlds parameter. This special case corresponds roughly to the Cratylan view that a proper name's referent is essential to it.

Let us now consider how to define "semantic neutrality." One might propose (5)'s negation as a definition. However, this proposal misses the thought that a type's essential nature leaves its semantics unconstrained. Imagine a type *t* such that

$$\exists d \exists e[d \neq e\ \&\ \Box\ \forall x(x \text{ is a token of type } t \rightarrow val(x, d) \vee val(x, e))].$$

Then *t* is not semantically permeated, yet neither does it sound right to say that *t* is semantically neutral. Intuitively, *t*'s underlying nature constrains it to assume one of two semantic values. We want to rule out such semantic constraints. I propose that type *t* is semantically neutral iff

[8] D. Lewis, *Philosophical Papers,* vol. 1 (Oxford UP, 1983).

(6) $\Box\, \forall d\, [\Diamond \exists w \exists z(z \text{ is a token of type } w \;\&\; val(z, d)) \rightarrow \Diamond \exists z(z \text{ is a token of type } t \;\&\; val(z, d))]$.

Intuitively, *t*'s tokens can assume any semantic value that is an admissible semantic value. There are various maneuvers through which one could then capture the intuitive idea that *z* might have had no meaning. I introduce the following convention: if *z* has no meaning, then $val(z, \perp)$.

(5) and (6) eschew some familiar notions from historical and contemporary discussions. Beginning with the *Cratylus*, the question of whether natural language words are semantically neutral has often accompanied the question of whether they refer due to *social convention*. (5) and (6) do not mention social convention. One could combine the claim that words satisfy (6) with the view, espoused by Davidson, that convention does not underpin linguistic meaning. To take another example, many philosophers who favor semantically permeated concepts maintain that concepts "resemble" what they represent. This thesis is optional for those who gloss semantic permeation through (5).

The definitions (5) and (6) are highly general. They encompass all three domains mentioned in §1: natural language, mental representation, and formal logic. For the rest of the paper, I will focus on the special case of *mental representation*. I will explore how the conflict between semantically neutral and semantically permeated approaches arises within contemporary discussion of Mentalese. I will then deploy my modal definitions to argue that the semantically permeated viewpoint has far more potential than the current literature recognizes.

§3. The representational theory of mind

A familiar view posits *mental representations* as the basic building blocks of perception and thought. These entities are analogous to concrete sentences, maps, diagrams, and pictures, but they are "lodged in the mind" rather than in the external world. More precisely, define *the representational theory of mind* (RTM) by the following three doctrines:

> Any mental state or event with representational properties is a relation to a mental representation.
>
> The representational properties of a mental state or event are inherited from the representational properties of the mental representation to which it is a relation.
>
> Psychological processes operate over mental representations.

All three doctrines have been widely discussed in the contemporary literature. I presume they are familiar enough not to require extended elucidation here.

As I have formulated RTM, it takes no sides regarding the central issue of the paper: semantic neutrality versus semantic permeation. RTM claims that representationally significant mental representations serve as the vehicles of mental activity. It does not say whether those vehicles are individuated partly by their representational significance. It is consistent with insisting either that mental representations are semantically permeated or that they are semantically neutral, in the sense of definitions (5) and (6). It is also consistent with a hybrid position that accepts mental representations of both kinds.

To illustrate, consider the belief that *that dogs are friendly*. On typical modern versions of RTM, this belief is a relation some Mentalese sentence DOGS ARE FRIENDLY. Let us suppose that the sentence contains a predicate DOG denoting dogs and another predicate FRIENDLY denoting friendliness. On a semantically neutral approach, DOG could just as easily have denoted cats, and FRIENDLY could just as easily have denoted moodiness. DOGS ARE FRIENDLY could just as easily have meant *that cats are moody*, rather than *that dogs are friendly*. DOGS ARE FRIENDLY is a piece of formal syntax, lacking any inherent interpretation. In contrast, the semantically permeated approach posits no such level of formal syntax. DOG denotes dogs, by its essential nature. FRIENDLY denotes friendliness, by its essential nature. DOGS ARE FRIENDLY means *that dogs are friendly*, by its essential nature. It is not subject to reinterpretation as meaning *that cats are moody*.

Historical philosophers tend to assume the semantically permeated viewpoint. Ockham, an early champion of Mentalese, endorses a semantically permeated conception of it. He regards propositional attitudes as relations to sentences in a "mental language" with *natural signification*, as opposed to the *conventional signification* of spoken and written language. Mentalese types come with their contents attached. Ockham does not endorse a level of non-semantic description. There is no suggestion in his writings that we can "hive off" a mental representation's content, leaving behind an explanatorily significant non-semantic residue.

Modern versions of RTM are typically far more hospitable towards semantic neutrality. The evolution of Fodor's views provides a good illustration.

Fodor's early view hinges on the distinction between narrow versus wide content.[9] The former kind of content supervenes on the thinker's internal neurophysiology, while the latter does not. Fodor's early view holds that Mentalese types have narrow contents essentially and that narrow content does not determine reference or truth-conditions. Fodor's early view treats Mentalese words as permeated by narrow content but not by denotation or wide content. Technically, this view embraces semantic permeation. However, Fodor's narrow contents are so detached from familiar referential semantics that for many purposes one might as well embrace semantic neutrality. Fodor grows increasingly hostile over time to semantic permeation, although it still makes an occasional appearance in his writings.[10] By the mid-1990s, Fodor officially

---

[9] Fodor, *Representations* (MIT Press, 1981).
[10] Fodor, *A Theory of Content and Other Essays* (MIT Press, 1990), at p. 167.

abandons both narrow content and semantic permeation. On his mature view, content is a wide, and Mentalese lacks any inherent semantic import.[11] A Mentalese type can acquire content through appropriate causal or counterfactual links to the external world. But the same type could have had a completely different content if the causal or counterfactual links had been different.

Ultimately, then, Fodor embraces a semantic neutral view of Mentalese. On Fodor's mature view, Mentalese syntactic types conform to (6).

Fodor also countenances semantically permeated entities, which he calls "concepts."[12] These are individuated by two orthogonal parameters: semantically neutral Mentalese syntactic type; and denotation. The first parameter is a "vehicle" for, or a "mode of presentation" (MOP) of, the second. As he puts it, "a concept is a MOP together with a content" (p. 20), so that concepts "have their contents essentially" (p. 120). Fodorian concepts conform to (5).

The key point here is that, although Fodor's mature view posits semantically permeated concepts, it *also* isolates an explanatorily crucial level of non-semantic individuation. As he puts it, "mental representations can differ in content without differing in their intrinsic, formal, nonrelational, nonsemantic properties… [T]he properties of mental states in virtue of which they are engaged by mental processes are intrinsic/syntactic."[13] Fodor recognizes semantically permeated mental entities, but he insists that any such entity factors into a formal, syntactic, non-semantic component (a "vehicle" or an "MOP") and an arbitrarily associated denotation. Only the first component determines a representation's role in mental activity: "it is the syntax, rather than the content, of a mental state that determines its causal powers."[14] Our best scientific psychology should taxonomize mental states through semantically neutral types, and it should

---

[11] Fodor, *The Elm and the Expert* (MIT Press, 1994), pp. 22-24; *LOT2*, pp. 69-80.
[12] Fodor, *Concepts* (Oxford UP, 1998).
[13] Fodor, "Replies," in B. Loewer and G. Rey (eds), *Meaning in Mind*, (Cambridge: Blackwell, 1991), p. 298.
[14] Fodor, *LOT2*, p. 70.

isolate mental processes defined over these types. Virtually all current philosophical advocates of Mentalese follow Fodor in emphasizing a semantically neutral level of syntactic individuation.

In my view, the current emphasis on semantically neutral Mentalese types is mistaken. I favor a pendulum swing towards the old-fashioned Ockhamite emphasis upon semantically permeated Mentalese types. I think that many (although perhaps not all) mental representations have fixed meanings, by their essential natures. No doubt one *can* adopt a taxonomic scheme that type-identifies these mental representations in semantically neutral fashion. I do not deny this possibility. I merely suggest that, in many cases, such a taxonomic scheme contributes no value to theorizing about the mind. It introduces semantically neutral types that serve no important explanatory purpose. One gains nothing by "factoring" the semantically permeated mental representation into a semantically neutral component and an arbitrarily associated content.

My goal is to explore the prospects for this revived Ockhamite viewpoint.

§4. Ontology, individuation, and explanation

Opponents of semantically permeated mental representation often ridicule it as mysterious or obscure. Putnam complains that "[n]one of the methods of representation we know about has the property that the representations *intrinsically* refer to whatever it is that they are used to refer to" (p. 21).[15] He warns us to be "highly suspicious of theories that postulate a realm of 'representations' with *such* unlikely properties" (p. 22). He implies, without asserting, that positing such representations is tantamount to positing entities with "magical" powers. In response, Burge denies that there is anything mysterious about semantically permeated mental

[15] H. Putnam, *Representation and Reality* (MIT Press: 1988).

items. As he puts it, "I see no reason to construe an explanatory scheme that identifies mental abilities in terms of their specific intentional functions (concepts) as 'magical'."[16]

I agree with Burge. A key point here is that semantically permeated types are *types*, whose primary role in our discourse is to taxonomize tokens. Any systematic psychological theory must adopt a suitable taxonomic scheme for classifying token mental states and events. A collection of types reifies a taxonomic scheme, providing each relevant category of tokens with an "ontological correlate." The types are abstract entities corresponding to our classificatory procedures. Semantically neutral types correspond to a taxonomic scheme that does not take meaning or content into consideration. Semantically permeated types correspond to a taxonomic scheme that classifies tokens, at least in part, by their meanings or contents.

Reifying a taxonomic scheme is a substantive move. Ultimately, we should ground this move in a rigorous modal theory of abstract entities, including paradox-free comprehension axioms attesting that types with desired modal properties exist. For present purposes, I assume that any reasonable taxonomic scheme corresponds to a suitable range of existing abstract types.

Nominalists, who reject abstract entities in general and linguistic types in particular, will object. As many philosophers have noted, however, quotidian and scientific discourse freely posit linguistic types. Nominalism poses no special challenge for semantically permeated types, because nominalists already face the more general task of paraphrasing or explaining away widespread apparent reference to abstract entities. Setting aside more general skepticism about abstract entities, there is no special *metaphysical* problem about semantically permeated entities. For instance, there is nothing "magical" or "mysterious" about a Mentalese word that refers to dogs by its essential nature. Semantically permeated types are mere "ontological correlates" to a

---

[16] Burge, *Foundations of Mind*, p. 302.

taxonomic scheme that type-identifies states, events, or processes partly through their meanings. The real question is whether such a taxonomic scheme supports a satisfactory theory of the mind.

Thus, the central issue here is not ontological but explanatory. How do our best scientific explanations type-identify mental states, events, and processes? In particular, do our best scientific explanations of any mental phenomena require a semantically neutral individuative scheme for mental representations?

Fodor and his followers frequently motivate a semantically neutral factor by citing current scientific practice. They hold that current scientific explanations feature inherently meaningless mental representations. Although I will not pursue the point here, I believe that this assessment distorts contemporary scientific practice. Many established branches of cognitive science feature no overt appeal to a formal, syntactic, non-semantic level of individuation. For instance, Burge forcefully argues that vision science individuates psychological states in terms of representational content (e.g. representing *that* an object is a certain distance from the perceiver), not in terms of any formal, non-semantic properties.[17] Vision science taxonomizes mental states through their representational properties, not through "formal syntax." Thus, if we want to apply RTM to vision, the natural strategy is to posit semantically permeated rather than semantically neutral representations. We can posit a mental representation that inherently represents a certain color, another mental representation that inherently represents a certain location in egocentric space, and so on. Nothing about vision science suggests any need to isolate a semantically neutral "factor" associated with each such representation.

As this example illustrates, many established branches of cognitive science do not match the template delineated by Fodor. The relevant explanations do not cite semantically neutral

[17] Burge, *Origins of Objectivity* (Oxford UP, 2010), at p. 96.

syntactic items. Perhaps *some* areas of cognitive science posit a semantically neutral syntactic factor. But many established areas do not.

Neuroscience type-identifies mental states by neural properties, which may lack any essential ties to semantic properties. However, this observation does not support semantically neutral versions of RTM. As Fodor notes, multiple realizability entails that neurally heterogeneous creatures can instantiate the same Mentalese syntactic type.[18] Thus, RTM should not individuate mental representations by their neural properties. Our question is whether RTM should employ a "formal syntactic" taxonomic scheme orthogonal both to neural taxonomization *and* to taxonomization by meanings.

Setting aside current scientific practice, philosophers offer various additional arguments that scientific psychology requires semantically neutral mental representations. Some of the most prominent arguments cite: physicalism; the problem of intentionality; the demand that scientific explanations be sufficiently general (often coupled with an appeal to Twin Earth); the obscure individuation conditions of mental content; worries about intentional causation (how can semantically permeated mental representations be causally efficacious?); and the claim that one can model the mind as a computational system only if one posits semantically neutral vehicles of computation. I address some of these arguments in (Author's A).

For purposes of this paper, I set such arguments aside. Instead, I will address a particularly vexing challenge that faces proponents of semantically permeated mental representation: how to accommodate *syntactic structure*. Most historical and contemporary versions of RTM posit "structured" mental representations. These are constructed from simpler representations by applying "compounding devices." In Fodor's words, then, "[w]hat we're all

[18] Fodor, *LOT2*, pp. 90-91.

doing is really a kind of logical syntax (only psychologized)."[19] Talk about "syntax" naturally suggests a contrast with "semantics," thereby indicating a role for inherently meaningless entities. Don't we require semantically neutral types to serve as syntactic structures? How can an Ockhamite approach accommodate the crucial thought that mental representations are syntactically structured? I will address these worries by studying "syntax," as it figures in both semantically neutral and semantically permeated theories.

§5. Syntactic compounding devices

I begin by exploring the syntactic compounding devices employed by modern logicians. Logic textbooks often construct formal languages from a finite alphabet of primitive symbols type-identified by their geometric forms. The textbook lists various tokens, such as

( ) *x* *y* & ¬

These tokens have geometric shapes that individuate primitive types composing the alphabet. Complex types are generated from primitives type by physical concatenation, which places symbols in physical proximity to one another. An inductive definition specifies which linear sequences of concatenated symbols count as "well-formed formulas."

Some logicians adopt a more abstract approach to formal languages. "For metamathematical considerations," Gödel notes, "it does not matter what objects are chosen as primitive signs."[20] Thus, primitive elements need not be individuated by geometric shape. For mathematical purposes, they could be anything at all. Similarly, the fact that concatenation places symbols adjacent to one another in *physical* space does not matter for purposes of proving mathematical theorems about formal languages. What matters is that concatenation conforms to

[19] Fodor, *Representations*, p. 223.
[20] K. Gödel, "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I," repr. in his *Collected Works,* vol.1, S. Feferman, et al. (eds), (Oxford UP, 1986), p. 147.

certain structural axioms (e.g. concatenating the null string with another string *x* yields *x* as a result). These observations suggest a purely "axiomatic" approach to formal languages, on which a formal language is built from arbitrary primitive objects through an operation ("concatenation") that satisfies certain general axioms, such as Tarski's.[21] On this approach, the essential nature of concatenation is exhausted by some axiomatization that makes no reference to parochial factors such as spatial proximity.[22]

Thus, the logical tradition offers at least two approaches to "syntactic compounding devices." On the first approach, the basic compounding device is *physical concatenation*, which has constitutive connections to proximity in physical space. On the second approach, the basic compounding device is *axiomatic concatenation*, whose inherent nature is exhausted by relevant axioms, such as Tarski's. A syntactically structured linguistic expression results from iteratively applying one of these two compounding devices to an alphabet of primitive symbols.

Given this background, let us now ask what notion of "syntactic structure" applies to mental representation. Most modern versions of RTM emphasize concatenation, without explicitly distinguishing between physical or axiomatic construals. C. R. Gallistel and Adam King offer a typically ambiguous, albeit relatively nuanced, discussion. They take as paradigmatic the combinatorial syntax that arises in a digital computer, where "the constituent symbols are atomic data (bits) that are physically adjacent in a strict ordering."[23] This quotation suggests physical concatenation as the compounding device. In a footnote, however, Gallistel and King admit that "[t]echnically, the data need not be physically adjacent. The architecture of

---

[21] A. Tarski, "The Concept of Truth in Formalized Languages," repr. in his *Logic, Semantics, Meta-Mathematics*, 2nd ed., J. H. Woodger (trans), J. Corcoran (ed), (Indianapolis: Hackett, 1983), pp. 152-278.
[22] Cf. C. Parsons, *Mathematical Thought and its Objects*, (Cambridge UP, 2008), p. 52.
[23] C. R. Gallistel and A. King, *Memory and the Computational Brain*, (Malden: Wiley-Blackwell, 2009), p. 150.

the machine could realize such a *next to* operation using some arbitrary (but fixed in the architecture) topological relationship." The footnote suggests an axiomatic conception.

For our purposes, the difference between physical and axiomatic concatenation is relatively unimportant. The key point is that *both* compounding devices are semantically neutral. Neither compounding device has any inherent semantic import. For instance, simply placing two symbols next to one another in physical space has no inherent meaning. It has only meaning only once we confer meaning upon it, either by explicit stipulation or by tacit convention. A similar point applies to axiomatic concatenation. A fixed relation that satisfies Tarski's axioms does not by virtue of that fact display any particular semantic import. As far as the axioms go, it might have whatever semantic import we want, or no semantic import at all. Both physical and axiomatic concatenation are inherently meaningless syntactic operations.

For that reason, neither device is suitable for a semantically permeated approach to mental representation. Complex mental representations, not just primitive representations, should have inherent meaning. To achieve that end, the complex representations must arise through syntactic operations with inherent compositional significance. Otherwise, we will simply have inherently meaningful parts combined into an inherently meaningless syntactic structure.

I conclude that a satisfactory semantically permeated version of RTM should assume the following form. We posit a "mental language" composed of finitely many primitive types. Each primitive type is semantically permeated. Complex expressions are generated by "semantically permeated syntactic operations": modes of syntactic combination individuated partly by semantic significance. These are devices for compounding types into more complex types, each such device carrying inherent semantic import. A complex type decomposes into semantically permeated parts arranged in an inherently meaningful syntactic configuration. The inherent natures of the parts and the inherent nature of the syntactic configuration jointly ensure that the complex type is permeated by an appropriate semantic value. Ockham already articulated essentially this picture of Mentalese syntactic structure.[24]

What exactly is a "semantically permeated syntactic operation"? The suggestion that syntactic operations can carry inherent semantic import may appall those readers accustomed to inherently meaningless operations, such as physical or axiomatic concatenation. Higginbotham rejects any such approach: "the combinatorial rule that applies to a structure cannot itself be determined by the structure to which it applies. In this sense, instructions for interpretation must be added to the bare syntactic structure."[25] Higginbotham does not defend this conclusion. But his remarks reflect a standard approach within modern logic, according to which a formal language consists of formulas that lack inherent meaning and hence admit alternative interpretations. Logicians may associate a formal language with an "intended interpretation," but the interpretation does not inhere in the language itself. This approach leaves little room for semantic permeation, let alone for inherently meaningful modes of syntactic combination.

I think that the notion of a "semantically permeated syntactic operation" is far more compelling than Higginbotham's remarks suggest. Something like this notion naturally arises *at the level of content* as a plausible reaction to the "unity of proposition." Suppose we associate sentences with "structured meanings," such as Fregean thoughts or Russellian propositions. A complex meaning is "compounded" from simpler meanings. But what binds those simpler meanings together into a unified whole? If the "compounding operation" is exhausted by a set-theoretic construction (such as an ordered pair or a tree-structure), then we do not yet obtain a

---

[24] P. King, "*William of Ockham: Summa Logicae*," in J. Shand (ed), *Central Works of Philosophy*, vol. 1, (McGill-Queen's UP), pp. 242-270.
[25] J. Higginbotham, "Expression, Truth, Predication, and Context: Two Perspectives," *International Journal of Philosophical Studies* (16), pp. 473-494, at p. 492, fn. 1.

genuinely meaningful complex. A set-theoretic construction can be interpreted in many different ways. For instance, if we combine the sense of "Theaetetus" and the sense of "flies" into the ordered pair <*sense*("Theaetetus"), *sense*("flies")>, we do not yet attain a compound thought that represents Theaetetus as flying. The ordered pair, in itself, could just as well be interpreted as meaning that Theaetetus does not fly. As Jeff King notes, a natural solution is to build "structured meanings" through compounding devices that themselves have inherent compositional import.[26] King develops this solution within a broadly Russellian framework. I believe that something like this solution is implicit in Frege's discussion of "thoughts." The "structure" of a Fregean thought carries rules for its own interpretation.

Our main concern here is not *meanings* but *vehicles* of meaning. Our concern is the linguistic types (especially Mentalese types) that express meanings. Can we make sense of a complex linguistic type whose internal structure carries inherent semantic import? Once again, Frege provides a useful model. As I will now argue, Frege bases his philosophical logic on semantically permeated syntactic operations for generating complex *Begriffsschrift* expressions.

§5.1 Frege on semantic permeation

Frege distinguishes between a *figure* (*Figur*), which is an "artifact consisting, perhaps, of printer's ink" and whose properties are "geometric, physical and chemical ones," and a *sign* (*Zeichen*), which has the "purpose of designating" as part of its essential nature (pp. 97-8).[27] A sign results from imbuing a figure with a sense. Thus, a sign comes with its interpretation attached. A *sentence* (*Satz*) is "a group of signs that expresses a thought" (p. 377), that is, the result of combining signs together in a meaningful combination. Were we to alter the thought

---

[26] J. King, *The Nature and Structure of Content* (Oxford UP, 2007).
[27] G. Frege, *On the Foundations of Geometry and Formal Theories of Arithmetic*, E.-H. Kluge (trans), (New York: Yale UP, 1971).

expressed by a sentence, "it would not even be the same sentence; not, at least, if one considers the thought expressed in it essential to the sentence" (p. 323). Thus, as May and Antonelli emphasize, Fregean languages have inherent interpretations.[28] In my terminology, elements of a Fregean formal language are semantically permeated.

Famously, Frege treats compositionality as function application. He models predicates and logical connectives as function-expressions, and he assimilates their linguistic behavior to a more general theory of functions. His theory rests on the primitive syntactic operation *inserting a singular term into a function-expression's argument-place*. By doing so, one "saturates" an "unsaturated" expression. I will call this syntactic operation *functional insertion*.

Frege does not regard functional insertion as requiring further semantic interpretation. The picture is not this: a function-expression denotes a function; now we combine it with a singular term, which denotes an object; and finally we stipulate, as an additional step, what the resulting complex expression denotes. A function-expression has argument-places as part of its essential nature: "every function sign must always carry with it one or more places which are to be taken by argument signs; and these argument places… are a necessary component of the function sign."[29] In that sense, the function-expression taken by itself is inherently "incomplete." The basic syntactic device is completing a function-expression by filling its argument-places: "we obtain a name of a value of a function for an argument, if we fill the argument-places in the name of the function with the name of the argument. In this way, for example, '$(2+3\cdot1^2)\cdot1$' is a name of the number 5, composed of the function name '$(2+3\xi^2)\cdot\xi$' and '1'."[30] The primitive

[28] R. May and A. Antonelli, "Frege's New Science," *The Notre Dame Journal of Formal Logic* 41 (2000), pp. 242-270, at p. 245.
[29] Frege, *Philosophical and Mathematical Correspondence*, (eds) G. Gabriel, et al., (trans) H. Kaal, (Oxford: Blackwell, 1980), p. 116.
[30] Frege, *The Basic Laws of Arithmetic*: *Exposition of the System*, (ed and trans) M. Furth, (Berkeley: University of California Press, 1964), p. 34.

syntactic device in Frege's system is an inherently meaningful mode of combination, not a meaningless operation such as physical or axiomatic concatenation.

Frege acknowledges that figures do not have their argument-places essentially: "[a]s a mere thing, of course, the group of letters 'and' is no more unsaturated than any other thing. It may be called unsaturated in respect of its employment as a sign meant to express a sense…: its purpose as a sign requires completion by a preceding and completing sentence."[31] By imbuing a figure (the mere "group of letters 'and'") with an appropriate sense, we thereby generate an unsaturated function-expression denoting the appropriate truth-function. Essential to the function-expression's nature is the potential for us to fill its argument-places with singular terms (paradigmatically, sentences denoting truth-values), thereby yielding a complex expression denoting an appropriate truth-value.

Frege's discussion is not an ideal paradigm for Ockhamite approaches to the mind, for three reasons. First, generations of philosophers have condemned Frege's talk about "unsaturated" function-expressions as unhelpful or even hopelessly obscure. Second, Frege says virtually nothing about modality, which was crucial to my analysis in §2. Third, a Fregean *Zeichen* factors into an semantically neutral *Figur* and an arbitrarily associated sense. I now offer a toy example that avoids these three worries while remaining broadly Fregean in spirit.

§5.2 A semantically permeated language

I work within Lewis's intensional semantics. A singular term intension is a function from indices to things. A sentence intension is a function from indices to truth-values. A predicate intension is a function from singular term intensions to sentence intensions: i.e., from functions

[31] Frege, "Compound Thoughts," repr. in his *Collected Papers*, P. Geach and R. Stoothoff (trans), B. McGuinness (ed), (New York: Blackwell, 1984), pp. 390-406, at p. 393.

from indices to things to functions from indices to truth-values. The basic compositional device in Lewis's system, as in Frege's, is function application. For instance, the meaning of an atomic sentence "*Fa*" is determined compositionally by the following clause:

(7) The intension of "*Fa*" is the result of applying the intension of "*F*" to that of "*a*."

Lewis develops his intensional semantics for a semantically neutral language, whose syntactic structures are given by inherently meaningless phrase structure trees. As I will now illustrate, one can easily expand Lewis's approach to include semantically permeated languages.

I begin with a simple language of semantically permeated atomic sentences. The language contains finitely many singular terms

$\mathbb{s}_1, \ldots, \mathbb{s}_m$

and finitely many one-place predicates

$\mathbb{F}_1, \ldots, \mathbb{F}_n$

Each singular terms $\mathbb{s}_i$ expresses, as part of its essential nature, an appropriate intension ϕ_i:

(8) $\Box\, \forall x(x$ is a token of $\mathbb{s}_i \rightarrow val(x, \phi_i))$

Each predicate $\mathbb{F}_i$ expresses, as part of its essential nature, an appropriate intension φ_i:

(9) $\Box\, \forall x(x$ is a token of $\mathbb{F}_i \rightarrow val(x, \varphi_i))$

Thus, $\mathbb{s}_i$ and $\mathbb{F}_i$ are semantically permeated.[32] An atomic sentence $\mathbb{F}_i(\mathbb{s}_k)$ results from applying functional insertion to $\mathbb{s}_k$ and $\mathbb{F}_i$. More rigorously, we take as a primitive notion "*x* results from inserting token *y* into an argument-place of token *z*," and we let $\mathbb{F}_i(\mathbb{s}_k)$ be a new type such that:

(10) $\Box\, \forall x(x$ is a token of $\mathbb{F}_i(\mathbb{s}_k) \leftrightarrow \exists y \exists z(y$ is a token of $\mathbb{F}_i$ & z is a token of $\mathbb{s}_k$ & x results from inserting z into the argument-place of $y))$

[32] To avoid any threat of paradox or circular membership-chains, I assume a domain of "things" that includes neither intensions nor semantically permeated types. Intensions are defined over this domain of things. (Cf. Lewis, *Philosophical Papers,* vol. 1, pp. 196-197.)

By its essential nature, functional insertion conforms to (7). More rigorously:

(11) $\Box$ [$val(y, \varphi_i)$ & $val(z, \phi_k)$ & x results from inserting z into the argument-place of $y \rightarrow val(x, \varphi_i(\phi_k))$]

From (8)-(11), we can derive within any reasonable quantified modal logic that

(12) $\Box\, \forall x[x$ is a token of $\mathbb{F}_i(\mathbb{s}_k) \rightarrow val(x, \varphi_i(\phi_k))]$

so that each atomic sentence $\mathbb{F}_i(\mathbb{s}_k)$ is permeated by the desired intension.[33] Moreover, each atomic sentence type is composed of semantically permeated primitive types ($\mathbb{F}_i$ and $\mathbb{s}_k$) combined through an inherently meaningful syntactic device (*functional insertion*). In a natural sense, then, each atomic sentence comes with its compositional semantics "built in."

We can easily extend our construction to logically complex expressions. For instance, we can introduce a new primitive type $\mathbb{\&}$ such that

(13) $\Box\, \forall x(x$ is a token of $\mathbb{\&} \rightarrow val(x, conj))$

where *conj* is an intension corresponding to conjunction. More precisely, *conj* is a two-place function that takes as input two functions from indices to truth-values and yields as output a function from indices to truth-values; the output function maps an index to *truth* just in case each input function maps that same index to *truth*. We can then inductively define $\mathbb{S}\ \mathbb{\&}\ \mathbb{T}$ as the result of inserting $\mathbb{S}$ and $\mathbb{T}$ into $\mathbb{\&}$'s two argument-places. (Here, we can either modify our treatment of functional insertion so as to handle multi-place functions, or else we can "curry" $\mathbb{\&}$ so as to convert it into a one-place function.) One can similarly introduce semantically permeated negation operators $\neg$ and $\vee$. One can then define an infinite language $\mathbb{L}$ of semantically permeated types through a standard inductive definition:

(14) $\mathbb{F}_i(\mathbb{s}_k)$ is a sentence of $\mathbb{L}$, for any predicate $\mathbb{F}_i$ and any singular term $\mathbb{s}_k$

[33] I assume that functional application is rigid: $y = f(x) \rightarrow \Box\, y = f(x)$. This assumption ensures that we can quantify into (12), thereby deriving that $\mathbb{F}_i(\mathbb{s}_k)$ is semantically permeated.

(15) $\mathbb{S} \,\&\, \mathbb{T}$ is a sentence of $\mathbb{L}$, for any sentences $\mathbb{S}$ and $\mathbb{T}$ of $\mathbb{L}$

(16) $\mathbb{S} \vee \mathbb{T}$ is a sentence of $\mathbb{L}$, for any sentences $\mathbb{S}$ and $\mathbb{T}$ of $\mathbb{L}$

(17) $\neg\mathbb{S}$ is a sentence of $\mathbb{L}$, for any sentence $\mathbb{S}$ of $\mathbb{L}$

(18) A type is a sentence of $\mathbb{L}$ only if its being so follows from (14)-(17)

Through an inductive proof, we can show within any reasonable quantified modal logic that each element of $\mathbb{L}$ is permeated by the desired intension. We could extend the construction to handle quantification. However, quantification raises complexities irrelevant to my discussion here.

Each element of $\mathbb{L}$ has a "configuration-structure-type" independent of the particular primitive elements composing it. For instance, the two sentences

$\mathbb{F}_1(\mathbb{s}_4) \,\&\, \mathbb{F}_5(\mathbb{s}_2)$

$\mathbb{F}_3(\mathbb{s}_5) \vee \mathbb{F}_2(\mathbb{s}_6)$

have the same configuration-structure-type, because they apply functional insertion in the same pattern to different primitive elements. In a natural sense, the two sentences share the same "syntactic structure." Tokens of a configuration-structure-type exhibit uniform compositional import: function application, iterated in a particular way. Our construction demonstrates that, despite Higginbotham's skepticism, a syntactic structure can carry semantic import as part of its fundamental nature. A syntactic structure can determine rules for its own interpretation.

My discussion of $\mathbb{L}$ is schematic in several respects. I said nothing about the natures of $\mathbb{F}_i$, $\mathbb{s}_i$, or $\&$, beyond clauses (8), (9), and (13), respectively. Those clauses are consistent with a wide variety of conceptions. They allow that the natures of primitive types are *exhausted* by their respective intensions, so that a token has an appropriate type if it expresses the appropriate intension. Equally, they allow that primitive types are partly individuated by factors beyond their intensions, such as phonological or geometric properties, or modes of presentation, or functional roles in a cognitive or linguistic system, and so on. My discussion of "functional insertion" was similarly schematic. I assumed only that functional insertion conforms to (11). Does that assumption exhaust the nature of functional insertion? Or does functional insertion have further properties, such as properties associated with physical concatenation? My treatment is neutral.

My discussion is schematic in another important respect: I simply assume that appropriate types exist. As noted in §3, a complete account would integrate that assumption into a theory of abstract objects and their modal properties.

Nevertheless, my discussion shows that there is nothing mysterious about a syntactic compounding device that carries rules for its own interpretation. There is no tension between treating Mentalese as syntactically structured and treating it as semantically permeated. To develop such an Ockhamite view, we treat the mind as deploying ingredients of two kinds: primitive mental representations, which have inherent meanings; and primitive devices for combining representations into more complex structured representations, where the resulting structure itself carries inherent compositional import. We employ the resulting complex types to taxonomize mental states, events, and processes. Nothing about my construction of $\mathbb{L}$ suggests the need for an explanatorily prior semantically neutral "factor." We can posit a "syntactically structured language of thought" without positing a range of semantically neutral syntactic types.

The intensions that language $\mathbb{L}$ associates with predicates could just as easily be wide as narrow. Thus, my discussion illustrates that we can treat Mentalese types as syntactically structured even while holding that they are permeated by wide content. Such types reify a taxonomic scheme that classifies mental states partly through their wide contents.

§6. Conceptions of syntax

I have spoken about "semantically permeated syntactic types." But the very suggestion may strike some readers as oxymoronic. Don't we learn at the beginning of any decent logic class that syntax and semantics are distinct? How can syntax be implicitly semantic?

I admit that theorists sometimes define "syntax" so as to preclude semantic permeation. For instance, Church writes that "[t]he study of the purely formal part of a formalized language in abstraction from the interpretation… is called *syntax*."[34] The basic idea in Church's discussion, and in many others, is that syntax studies how inherently meaningless parts of certain types are arranged in inherently meaningless configurations of certain types. This conception, which regards syntactic structures as inherently meaningless, recurs frequently in linguistics and philosophy.[35] Let us call it ***the semantically neutral conception of syntax***.

However, there are at least two other conceptions of syntax that frequently appear in linguistics, computer science, and mathematical logic. I call these *the combinatorial conception* and *the finitary conception*. Neither conception entails the semantically neutral conception.

On ***the combinatorial conception***, syntax studies how parts of certain types are configured in structures of certain types. This conception, already suggested by etymology (*syn* = together, *tax* = arrangement), recurs frequently in the literature. In its most general form, the combinatorial conception takes no stand regarding the natures of part-types and configuration-structure-types. It treats syntax as the general study of how entities belonging to certain categories are combined into certain patterns. For instance, Fine proposes a discipline of *universal abstract syntax*: a "general theory of constituent structure," based on three primitive notions: *occurrence of*, *occurrence in*, and *substitution*.[36] As Fine emphasizes, his conception is

[34] A. Church, *Introduction to Mathematical Logic I*, (Princeton UP, 1956), p. 58.
[35] N. Chomsky, *Syntactic Structures* (Paris: Mouton, 1957), p. 5; J. Haugeland, *Artificial Intelligence: The Very Idea* (MIT Press, 1985), p. 100.
[36] K. Fine, *Modality and Tense* (Oxford UP, 2005), p. 74.

quite general, encompassing not only strings of concatenated signs, but also matrices, diagrams, and propositions. Sometimes, authors restrict the combinatorial conception to studying how *linguistic expressions* are configured into more complex linguistic expressions. Thus, Ludlow and Neale define "syntax" as "the study of the properties of expressions that distinguish them as members of linguistic categories, and 'well-formedness,' that is, the ways in which expressions belonging to these categories may be combined to form larger units."[37] Since Ludlow and Neale do not define "word," "sentence," or "linguistic expression," it is unclear how much the restriction to language really limits the combinatorial conception.

The semantically neutral conception results from the combinatorial conception when we supplement it with a restriction upon part-types and configuration-structure-types: that they be inherently meaningless. *The finitary conception* results when we supplement the combinatorial conception with an orthogonal restriction: that they be *finite*.

On ***the finitary conception***, syntax studies how parts of certain types are arranged in *finite*, *discrete* configurations of certain types. The finitary conception is narrower than the combinatorial conception. For instance, the combinatorial conception allows infinitely long formulas to count as "syntactically structured," while the finitary conception does not. Since the human mind is essentially finite, the finitary conception rather than the combinatorial conception is typically more apt for modeling linguistic and cognitive phenomena.

The most widely discussed finitary configurations are logical formulas generated through physical concatenation of primitive symbols individuated by their geometric shapes. However, there are at least three respects in which such examples are unrepresentative. First, linear sequences are just a special case of finitary structures. Non-linear configurations, such as the

---

[37] Ludlow and Neale, "Syntax," in E. Craig (ed), *The Routledge Encyclopedia of Philosophy* (London: Routledge, 1998), pp. 246-253, at p. 246.

two-dimensional structures employed by Frege's *Begriffsschrift* or the phrase structure trees employed in generative linguistics, can also be finite and discrete. Second, the finitary conception allows a syntactic entity to be composed of *anything*: the Taj Mahal, the number seven, the concept of justice, etc. Third, the finitary conception allows syntactic compounding devices other than physical concatenation, such as axiomatic concatenation.

Semantic permeation conflicts with neither the combinatorial nor the finitary conceptions of syntax. A finite, discrete configuration can satisfy (5). There is no conflict between an entity having its semantic value essentially and the entity having finite, discrete structure. Language $\mathbb{L}$ from §5.2 illustrates the point. Each expression of $\mathbb{L}$ is a finite, discrete configuration. For instance, the sentence $\mathbb{F}_1(\mathbb{s}_4)$ & $\mathbb{F}_5(\mathbb{s}_2)$ contains five elements ($\mathbb{F}_1$, $\mathbb{s}_4$, &, $\mathbb{F}_5$, and $\mathbb{s}_2$) arranged in a configurational structure that results from iterated application of functional insertion. More generally, each member of $\mathbb{L}$ results from finitely many iterations of a single syntactic operation (*functional insertion*) to primitive types drawn from a finite alphabet. Yet each element of $\mathbb{L}$ is semantically permeated, and each element contains semantically permeated parts.

I conclude that, at least on some legitimate and well-established uses of the word "syntax," syntactically structured items can be semantically permeated. Of course, the word "syntax" itself is not terribly important. Readers who insist on the semantically neutral conception of syntax are free to reinterpret all my talk about syntax as involving a new term "shyntax." These terminological issues should not distract us from the crucial point, which is the differences among the combinatorial, finitary, and semantically neutral conceptions.

In particular, a key moral is that *the finitary conception of syntax does not entail the semantically neutral conception*. An entity that has "syntactic structure" according to the former conception need not have "syntactic structure" according to the latter. A finite, discrete structure-type can carry inherent semantic import. If we provide an argument that some domain involves syntax *in the finitary sense*, we do not thereby argue that the domain involves syntax *in the semantically neutral sense*. To move directly from the finitary conception to the semantically neutral conception is fallacious. Roughly, the fallacy conflates distinct notions of the "formal." "Formal" might express a concern with linguistic structure: which primitive types compose an expression, and in which configurational patterns are the types arranged? Or "formal" might mean "non-semantic," in which case the term evokes a form/content dichotomy. A theory that is "formal" in the first sense need not be "formal" in the second sense.

Framed this starkly, the fallacy may seem glaring. But essentially this fallacy recurs frequently among proponents of semantic neutrality. Philosophers often defend the finitary (or combinatorial) conception but then directly slide, without any further argumentation, to the semantically neutral conception. For instance, I believe that Carnap repeatedly commits this fallacy in *The Logical Syntax of Language*. For present purposes, I will illustrate with a more relevant example drawn from contemporary philosophy of mind.

§7. Systematicity and productivity

As we have seen, Fodor holds that Mentalese expressions are *syntactically structured*. He contrasts his view with a more generic form of intentional realism, which claims only that mental states have semantic properties (such as truth-conditions). As he puts it, "the question we're arguing about isn't, then, whether mental states have a semantics. Roughly, it's whether they have a syntax."[38] Fodor glosses the claim that mental states have syntax as follows: "mental states --- and not just their propositional objects --- typically have constituent structure."[39]

[38] Fodor, *Psychosemantics* (MIT Press, 1987), p. 138.
[39] Fodor, *Psychosemantics*, p. 136.

Elsewhere, he clarifies that "[t]he constituency relation is a part/whole relation… More precisely… if C is a constituent of C*, then it is metaphysically necessary that for every tokening of C* there is corresponding tokening of C."[40] He also writes that "syntactic properties… are constituted entirely by what parts a representation has and how these parts are arranged."[41] These official pronouncements suggest the combinatorial conception of syntax. However, Fodor frequently emphasizes that the human mind has only "finite means" at its disposal, which suggests the narrower finitary conception.

Fodor offers two "armchair" abductive arguments for Mentalese, based respectively on *productivity* (thinkers can entertain a potential infinity of distinct thoughts, but they have only finite means at their disposal) and *systematicity* (there are systematic connections between which thoughts a thinker can entertain). An adequate theory should explain these two crucial phenomena. What explanation might we offer? As Fodor notes, we can treat propositional attitudes as relations to mental representations with a combinatorial syntax and a compositional semantics. To explain productivity, we iteratively apply Mentalese compositional mechanisms to an initial set of finitely many primitive words, thereby generating infinitely many Mentalese sentences graspable in principle by the thinker. We can similarly explain systematicity. If having a thought is "being related to a structured array of representations," then "presumably, to have the thought that John loves Mary is to ipso facto have access to the same representations, and the same representational structures, that you need to have the thought that Mary loves John."[42] Thus, Mentalese helps explain both productivity and systematicity.

---

[40] Fodor, "Connectionism and the Problem of Systematicity (Continued): Why Smolensky's Solution Still Doesn't Work," *Cognition* 62 (1997), pp. 109-119, at p. 111.
[41] Fodor, *The Mind Doesn't Work That Way* (MIT Press, 2000), p. 20.
[42] Fodor, *Psychosemantics*, p. 151.

Fodor's explanation of productivity and systematicity hinges on two key claims. First, Mentalese sentences have "constituency structure," so that a thinker who grasps a Mentalese sentence can "recombine" its parts to form a new sentence. Second, Mentalese semantics respects constituency structure, so that recombination of a sentence's parts yields a new sentence with desired semantic properties.

Both claims are consistent with semantic permeation. Elements of $\mathbb{L}$ have "constituency structure" in Fodor's sense. For instance, it is metaphysically necessary that for every tokening of $\mathbb{F}_1(\mathbb{s}_4)$ there is a corresponding tokening of $\mathbb{F}_1$ (under the virtually tautologous assumption that inserting a token of $\mathbb{s}_4$ into the argument-place of a token of $\mathbb{F}_1$ requires tokening $\mathbb{F}_1$). Moreover, the constituency structure is semantically relevant *by its inherent nature*. Thus, a semantically permeated Mentalese modeled after $\mathbb{L}$ has compositionally relevant constituency structure.

Accordingly, we can hypothesize that Mentalese involves inherently meaningful syntactic operations, such functional insertion. We can explain productivity through iterated application of these syntactic operations to finitely many semantically permeated primitive words, generating a potential infinity of inherently meaningful sentences graspable in principle by the thinker. We can explain systematicity by noting that a complex Mentalese sentence contains inherently meaningful parts arranged in an inherently meaningful syntactic configuration, so that a thinker able to entertain a semantically permeated sentence (such as JOHN LOVES MARY) has the representational resources needed to entertain appropriately related semantically permeated sentences (such as MARY LOVES JOHN). A semantically permeated view of Mentalese readily explains the desired phenomena, without any appeal to a non-semantic level of individuation.

Nevertheless, philosophers often move directly from systematicity and productivity to a semantically neutral version of Mentalese. In a typical recent discussion, Fodor urges that "constituent structure is required to account for the systematicity/productivity of thought," and hence that we should posit a "syntactically structured" language of thought.[43] He assumes without argument that the relevant kind of "syntax" is individuated in a purely non-semantic fashion. He does not consider an alternative approach that eschews non-semantic individuation. He simply assumes that a proper explanation posits semantically neutral syntactic types and inherently meaningless syntactic compounding operations.[44] The same oversight recurs frequently in Fodor's writings and the surrounding literature. For instance, Tetzlaff and Rey invoke systematicity to argue that mental activity is defined over "syntactically specified representations," where "representations are either simple or complex, the complex ones being composed by concatenation of the simple ones."[45] They assume without argument that syntactic structure is generated through iterated concatenation (they do not specify physical or axiomatic), rather than through an inherently meaningful operation such as functional insertion. A similar slide informs the exposition of Gallistel and King, who move directly from productivity to a semantically neutral medium of mental representations.

I submit that Fodor and his followers fallaciously conflate distinct conceptions of formality. Systematicity and productivity may show that mental states have "form," in the sense of combinatorial structure. They do not show that Mentalese has "form" in a sense opposing "content." We cannot move directly from the thesis that Mentalese has combinatorial structure, finitary or otherwise, to the thesis that Mentalese involves inherently meaningless items waiting to be imbued with content. Any such move involves a fallacious slide from the combinatorial or

---

[43] *LOT2*, p. 61.
[44] Fodor offers a second argument for semantic neutrality: namely, that the computational theory of mind requires it (pp. 61-62). I critique this second argument in (Author's A). My point here is that, despite what Fodor claims, productivity and systematicity provide no independent support for semantic neutrality.
[45] M. Tetzlaff and G. Rey, "Systematicity and Intentional Realism in Honeybee Navigation," in R. Lurz (ed), *Philosophy of Animals Minds* (Cambridge UP, 2009), pp. 72-88, at p. 73.

finitary conception of syntax to the semantically neutral conception. The kind of “mental syntax” required to explain systematicity and productivity need not be semantically neutral.[46]

§8. Semantically permeated mental representation

I have elucidated the distinction between semantic neutrality and permeation, as it arises in the study of language, mind, and logic. I have also developed a semantically permeated version of the crucial thesis that mental representations are *syntactically structured*. Drawing on these elements, I have highlighted a fallacious slide frequently committed by proponents of semantic neutrality. I have not defended or even formulated a detailed semantically permeated version of RTM. That task is far too large for a single paper. Nevertheless, it seems clear that semantically permeated approaches to mental representation deserve greater attention than they currently receive. Whether similar conclusions apply to natural language and formal logic is a question for another occasion.

---

[46] Cf. G. Evans, *The Varieties of Reference* (Oxford UP, 1982), pp. 100-101.