Relational Approaches to Frege's Puzzle¹

Aidan Gray University of Illinois at Chicago

Abstract

Frege's Puzzle is a fundamental challenge for accounts of mental and linguistic representation. This piece surveys a family of recent approaches to the puzzle which posit representational relations. I identify the central commitments of relational approaches, and present several arguments for them. I also distinguish two kinds of Relationism - Semantic Relationism and Formal Relationism - corresponding to two conceptions of representational relations. I briefly discuss the consequences of relational approaches for foundational questions about propositional attitudes, intentional explanation, and compositionality.

1 Frege's Puzzle²

The challenge posed by Frege's puzzle for accounts of representation begins with Frege's Insight.

Frege's Insight The cognitive significance of a sentence or propositional attitude is not determined by its referential content.

The referential content of a representation is the way it represents properties and relations as distributed over objects. Representations which portray the same objects as bearing the same properties, or standing in the same relations, share referential content. Frege noted that representations can share referential content but differ in cognitive significance. (1 a) and (1 b) represent the same man as being born in the same place, but differ in cognitive significance.

- (1) (a) Bob Dylan was born in Minnesota
 - (b) Robert Zimmerman was born in Minnesota

What it means to say that (1 a) and (1 b) differ in cognitive significance is partly what is at issue in responding to Frege's puzzle. For an initial idea, consider an example. Sally knows Dylan under both names, but doesn't realize that she knows one person under two names. She grew up with him in Minnesota and sometimes reminisces about her friend Bobby Zimmerman. She believes that Bobby was born in Wisconsin and moved to Minnesota at a young age. She knows of the famous musician Bob Dylan. She believes that he is from New York. Imagine contexts A and B: In A, someone asserts, and Sally accepts, (1 a); in B, someone asserts, and Sally accepts, (1 b). In B, she takes herself to have learned something about her old friend. This kindles a desire to reconnect. She calls her mother to find out what happened to Zimmerman. In A, she takes herself to learn something about the singer. She realizes that a Minnesotan won a Nobel prize. She feels a swell of pride while listening to Desire.

¹[forthcoming in *Philosophy Compass*] This paper was written with the support of the Institute for the Humanities at the University of Illinois at Chicago. I have benefited from discussion with Mahrad Almotahari, Rachel Goodman, Dave Hilbert, Michael Murez, Angel Pinillos, Francois Recanati, Will Small, and William Taschek. Special thanks to Richard Heck and Robin Jeshion.

²The literature on Frege's Puzzle is voluminous. The central texts from Frege are (1892; 1918). Beyond the works cited later, the discussions which shape the general approach taken here are: (*McDowell*, 1977), (*Perry*, 1977), (*Burge*, 1979), (*Dummett*, 1981), (*Evans*, 1982, 1985), (*Taschek*, 1992), (*Millikan*, 1997), (*Kremer*, 2010).

Sally formed different beliefs in the contexts; she did so in virtue of the different significance of the assertions. That the *beliefs* differ in significance is exhibited by the fact that they lead to different actions, different inferences, *etc.* Broadly speaking, they have different roles in rationalizing psychological explanation. The difference in significance in the *assertions* is located in the fact that Sally's understanding and accepting them led to different beliefs. Take this as an initial characterization of the potential for differences in cognitive significance between representations which share referential content.

Here's the challenge:

Frege's Challenge What, beyond the referential content of a sentence or propositional attitude, determines its cognitive significance?

Frege answered it by positing *senses*. Senses are aspects of mental and linguistic representations which capture the way an object is "presented" in that representation. Frege offered a two-level account of content: each element of referential content is accompanied by a sense. The same referential content can be presented by different senses and the cognitive significance of a representation is determined by its sense.

This characterizes sense in terms of its explanatory role. The difficulty lies in saying what senses are. Frege, himself, was not explicit about this question. When characterizing senses, he often reached for descriptions. On this model, the sense, for Sally, of "Robert Zimmerman" is my childhood best friend, while the sense of "Bob Dylan" is the famous NYC folk singer. Cognitive differences are explained in terms of the difference between referring to Dylan via these different descriptions.

2 Relationism

Relationism is a non-Fregean approach to Frege's puzzle. In this section, I will characterize Relationism in terms of two abstract theses: **Cognitive Significance as Coordination** and **Relationism about Coordination**. In section (4), I will describe two substantive versions of Relationism.

To understand the abstract characterization of Relationism, we need some background. First, coordination: The way that I'm using the term here, "coordination" is a theoretically-neutral term.³ At the intuitive level, the coordination of referential content in a body of representations captures how objects are represented "as the same". When Sally thinks of Dylan that he wrote Hurricane and of him that he wrote Mozambique she thinks of him in the same way in those thoughts. Those thoughts about Dylan are coordinated. When she thinks of Dylan that he grew up in the house next door she is not thinking of him as the same again. That thought is uncoordinated with the first two.

We must distinguish objects being represented as the same, from objects being represented to be the same (Fine, 2009, pg. 40). Identity statements - like (2) - represent objects as being the same, but do not, as such, represent objects as the same.

(2) Robert Zimmerman is Bob Dylan

³I adopt the term from (*Fine*, 2009). I am not using it here precisely as Fine uses it. Fine uses "coordination" in different ways, but usually to characterize some aspect of his *theoretical account* of what I'm calling coordination. As I'm using it here, it is theory-neutral. Below when I discuss Fine's account, I'll use the term "semantic coordination".

The substance of this distinction will come from an account of coordination; the intuitive idea is that representing objects to be the same is *asserting* their identity, representing them as the same is *presupposing* it.

Coordination partially determines the logical status of a body of representations. If representations are coordinated then they are in 'logical contact' in a way that exceeds merely sharing referential content. Coordination exists in different domains: within a person's attitudes at a time; between a person's attitudes across time; between attitudes of different people; between linguistic items; between attitudes and utterances, etc.

Particular relational theories focus on different kinds of coordination. It's open to us to be relationists about some kinds of coordination and not about others, or to offer different relationist account of different kinds of coordination. This variety in coordination makes it difficult to give a general characterization. It manifests itself in different ways in different domains. For now, characteristic examples must suffice.

Within an individual's attitudes at a time, coordination affects which inferences are licensed, whether the agent is irrational, etc. From the coordinated beliefs that Dylan wrote Hurricane and that Dylan wrote Mozambique, Sally is entitled to infer that the person who wrote Hurricane wrote Mozambique. In contrast, from her uncoordinated beliefs that Dylan won a Nobel Prize and that Zimmerman is a childhood friend she is not entitled to infer that her childhood friend won a Nobel Prize. Given her belief that Dylan is musical and her belief that Zimmerman is not musical are uncoordinated, she is not irrational; she doesn't believe an 'explicit' contradiction.

Coordination plays a similar role in language; it establishes logical connections. These connections partially determine when inferences are licensed, which sentences are contradictions, which are tautologies, etc. But coordination also partially determines whether a speaker understands an utterance. If terms are coordinated, a speaker who believes that they refer to different things fails to grasp their significance (Fine, 2009, pg. 40), (Pinillos, 2011). Contrast (3 a) with (3 b).

- (3) (a) When Dylan went electric, he was living in NYC
 - (b) When Dylan went electric, Zimmerman was living in NYC

Consider an interpretation of (3 a) in which 'he' is anaphoric on 'Dylan'. On this interpretation, 'Dylan' and 'he' in (3 a), and 'Dylan' and 'Zimmerman' in (3 b) share their referential content (they all refer to Dylan). But someone who believed that 'Dylan' and 'Zimmerman' referred to distinct individuals would not thereby have misunderstood (3 b) - that a competent speaker can fail to realize that names are coreferential is the heart of Frege's Puzzle. Someone who believed that 'Dylan' and 'he' referred to different people, would fail to understand (3 a). 'Dylan' and 'he' are coordinated in (3 a).

The attitudes of different agents can be coordinated. Sally's 'Dylan' beliefs are in rational contact with her peers' 'Dylan'-beliefs in a way that they are not in contact with her peers' 'Zimmerman'-beliefs. It's less obvious how to characterize the marks of inter-personal coordination. One idea is that there is a robust notion of agreement or disagreement which applies interpersonally only when referential content is interpersonally coordinated. For example, if Sally's friend has a belief that he would express with 'Zimmerman has no musical talent' he wouldn't count as disagreeing with Sally's belief that Dylan is a musical genius (consider: those attitudes couldn't form the basis of a productive rational exchange).⁴

The first relationist thesis is Cognitive Significance as Coordination (CSC).

⁴Two notes: First, this is not to deny that there is another sense of agreement/disagreement that tracks referential content. The point here is just that there is also a more fine-grained sense which tracks coordinated content. Second, the fact that beliefs would be expressed using the same name is neither a necessary nor sufficient

CSC Differences in cognitive significance between representations with the same referential content are explained by coordination.⁵

According to \mathbf{CSC} , when representations with the same referential content differ in cognitive significance, we explain the difference by appealing to how they are coordinated with other representations. Take the beliefs that Sally forms in response to (1a) and (1b). Relationists characterize the difference between them in terms of which other mental representations they are coordinated with. The upshot of accepting (1b), is a belief which represents Dylan as from Minnesota, and is coordinated with a number of pre-existing beliefs about him: e.g., that he grew up next door in Minnesota, that Sally lost touch with him, etc. The upshot of accepting (1a), is a belief with the same referential content but which is coordinated with different beliefs: e.g., that he is a famous folk singer, that he often gave prickly interviews, etc.

CSC holds that the cognitive difference between the beliefs - the different inferences and actions they lead to - is explained by the different way that the referential content is coordinated with Sally's attitudes. With respect to language, CSC holds that the difference in significance between (1 a) and (1 b) is captured in terms of how accepting those sentences leads to differentially coordinated mental states. So Relationism is committed an account of coordination among mental representations, among linguistic representations, but also between mental and linguistic representations (*Heck*, ms., pg. 37), (*Fine*, 2009, Chp. 4).

A cautionary note: **CSC** is too strong. No Relationist, as far as I'm aware, maintains that *every* cognitive difference between referentially-equivalent representations is explained by coordination. Heck (2012, pg. 161) acknowledges that the difference between a first-person thought and a third-person thought might be explained otherwise than by coordination. **CSC** should be read as applying to central cases that have been discussed in relation to Frege's Puzzle. At minimum it applies to non-indexical propositional attitudes and to sentences which differ only in the substitution of coreferential names. **CSC** represents a proposal for carving out a kind of cognitive significance, and a program for seeing how far this style of explanation can extend.

Relationism's decisive break with Fregeanism comes with **Relationism about Coordination** (**RC**).

RC The coordination of referential content is not determined by the sameness of (or resemblance between) intrinsic representational features. It is a representational relation. ⁶

Before explaining **RC**, an important clarification: I'm allowing myself to talk about coordination as holding between symbols (i.e. word occurrences, or mental vehicles) and as holding between elements of content (i.e. objects as they appear in the content of a sentence or attitude). But it's important to keep the two ideas distinct because different varieties of relationism see the relation between the two kinds of coordination differently. When I'm taking about coordination between symbols, I'll say that two 'occurrences' are coordinated; when I'm talking about content, I'll say that two 'elements' are coordinated.

A final note about how to understand the idea that two *elements* of referential content are coordinated. When I speak of an element of referential content, I'm not speaking about an

condition for coordination. But there is a non-accidental connection between sameness/difference of name and coordination. An account of inter-personal coordination should explain the role that sameness of name plays in establishing coordination without entailing that it is necessary or sufficient. See also note (9).

⁵See (Fodor, 1990), (Taschek, 1992), (Taylor, 2003), (Fine, 2009, esp. Chp. 2-3), (Heck, 2012, esp. section 2-3). ⁶(Taschek, 1995, pg. 83), (Fine, 2009, pg. 42), (Pinillos, 2011), (Heck, 2012, esp. pg 144-5), (Schroeter, 2012), (Pryor, 2016)

object simpliciter but an object as it appears in the content of a token representation. When we refer to Dylan as an element of content we are considering Dylan as he appears in a particular representation of him. So Dylan's two appearances in the content of "Zimmerman admires Dylan" will count as distinct elements of content; his appearance in Sally's belief that Dylan admires Leonard Cohen will be a distinct element from his appearance in Sam's belief that Dylan admires Cohen.

According to the Fregean account of coordination (FC), coordination is sameness of sense.

FC Occurrences o and o' are coordinated iff the sense o = the sense of o'

Frege posited features of sense which make it suitable for capturing coordination in different domains. That sense is what is "grasped" by a thinker means that sameness of sense is transparent to a thinker; thus if two thoughts involve the same sense, they are transparently presentations of the same object and are thus coordinated. Senses are "objective" - the same sense can be grasped by different individuals - thus making sense appropriate for capturing inter-personal coordination.

RC rejects this picture. Relationists hold that the fact that occurrences are coordinated is not explained in terms of how each occurrence, individually, presents its object. Instead, coordination is explained in terms of how occurrences, or elements of content, are *related* to each other. As far as their *intrinsic* representational properties are concerned, there is no difference between referentially equivalent representations. Referentially equivalent representations can only differ in their relational representational features.

Consider the difference between being soul-mates and being married. Being soul-mates is a matter of a match or fit between the properties of two people. If the personalities, experiences, etc of X and Y match, they are soul-mates. This is consistent with X and Y never having met or interacted. Contrast this with $being \ married$. No facts about X and Y's personalities, tastes, etc determines whether they are married. One has to consider how X and Y are related - in particular, whether they have entered into certain social or legal relations. Relationists hold that coordination is like marriage and unlike being a soul-mate. No representational facts about occurrences, considered in isolation from each other, determine facts about coordination.

An issue here is the nature of the intrinsic/extrinsic distinction. On a standard understanding, no representational features of a sign are intrinsic to it; on this understanding, RC would be trivially true. What relationists have in mind is that intrinsic representational features are those which can be stated without reference to another representation (Fine, 2009, pg. 22). But this is too weak, because it is vocabulary-dependent. A Fregean could introduce an intrinsic representational feature by defining it in terms of a relation - see the discussion of equivalence-class Fregeanism below - thereby turning a relational semantic feature into an intrinsic one. Relationists need a stronger version of the intrinsic/relational distinction, one which bears explanatory weight. But one that is not so strong that it counts all representational features as relational.

The above characterization of Relationism - in terms of **CSC** and **RC** - is abstract. Fleshing it out requires giving substance to relational representational features. There are two strategies. **Formal Relationists** explain coordination in terms of representationally relevant relations that hold between the vehicles of content. **Semantic Relationists** explain coordination in terms of semantic relations at the level of content.⁷ I examine these positions in section (4).

⁷A note about terminology. Tashek's (1995; 1998) work is clearly Relationist though he doesn't use the term. The term comes from Fine (2009) who describes his view as "Semantic Relationism". Pinillos (2011; 2015) identifies as a Relationist. The term "Formal Relationist" was coined in (*Almotahari*, 2013) to characterize the view in (*Heck*, 2012, 2014). Taylor (2003; 2010), Schroeter (2012), and Pryor can also be classified as relationists.

3 Arguments for Relationism

3.1 Pessimistic Induction on Fregean Accounts

Relationists cannot argue against every possible account of sense. But they point to a structural feature of Fregean accounts which suggests that they will fail. An account of sense determines an account of coordination via **FC**. The challenge for the Fregean is to characterize individuation-conditions for sense, thus explaining coordination (*Taschek*, 1998, pg. 330), (*Schroeter*, 2012, section 4).

After Kripke (1980), philosophers largely abandoned the idea that senses are descriptions. But a tradition - see (Dummett, 1981, especially chp. 5), (Evans, 1985), (Peacocke, 1983) - held onto the idea that the sense is connected to reference-determination. We could hold that occurrences are coordinated if their reference is determined in the same way (Evans, 1985, pg. 301-302). But it's doubtful that sameness of reference-determination is transparent to a thinker. It is possible for the reference of occurrences to be determined in the same way, but for a thinker to be unaware of this. So sameness of reference-determination cannot determine coordination (because coordination is transparent⁸).

Take the Kripkean approach to reference-determination: the reference of a name-occurrence is determined by the communicative chain leading back to a baptism. Is it transparent to a thinker when occurrences of a name are links in the same communicative chain? It seems not. Consider Kripke's (1979) Paderewski-case. Peter learns about the famous Polish musician/statesman in two conversations. He fails to realize that he is being told about the same person, so comes to believe that Paderewski, the statesman, and Paderewski, the musician, are different people. Peter's references to Paderewski are not all coordinated. He believes of Paderewski, under one way of thinking about him, that he is musical and under the other, that he is a famous statesman. But he is not licensed to infer that a famous statesman is musical.

Fregeans might respond by claiming that the methods of reference-determination for the two Paderewski-thoughts are distinct. They might, for example, point to the fact that they were acquired by Peter in different initial contexts. It is hard to see, though, how this will help. First, this fact itself need not be transparent to Peter (cf. (Fine, 2009, pg 37)). Second, we can imagine a version of the case where Peter learned about Paderewski in a single context - in which he was told that he was a musician and a statesman - but because of memory-failure comes to believe that there are two distinct individuals (cf. (Pryor, 2016, pg. 334)).

This provides a recipe for arguments against sense. It seems always possible, because of cognitive limitations, for a thinker to fail to realize of occurrences whose reference is determined in the same way, that their reference is determined in the same way (for more examples, see Fine's (2009, pg 130) discussion of explicit definition, and Heck's (ms., pg. 21) discussion of Peacocke's (1992) account of sense).

3.2 Indistinguishable Senses

Fine writes:

[L]et us imagine a universe which is completely symmetric around someone's center of vision. Whatever she sees to her left is and looks qualitatively identical to something

⁸This assumption is widely, but not universally, made about intra-personal coordination. Campbell (1987) argues that sameness of sense is not transparent. If this argument were accepted, the dialectic between the Fregean and the Relationist would change.

she sees on her right (not that she conceptualizes the two sides as "left" and "right" since that would introduce an asymmetry). She is now introduced to two identical twins, one to her left and the other to her right, and she simultaneously names each of them "Bruce"; using a left token of "Bruce" for the left twin and a right token of "Bruce" for the right twin. The two tokens of "Bruce" are then always used in tandem so as not to disturb the symmetry. [...] She can even assert the non-identity of the two Bruces by simultaneously uttering the one token of "Bruce" from the left side of her mouth, the other token from the right [.] (2009, pg 36.)

The thinker's two Bruce-thoughts are un-coordinated. It would be informative to be told there is only one Bruce. This challenges \mathbf{FC} because the two thoughts about Bruce are associated with the same non-relational representational features. The Fregean cannot point to any representational difference between them which would explain why they were un-coordinated. The thoughts are numerically but not qualitatively distinct (for critical discussion, see (Gray, 2016)).

But do the two thoughts about Bruce share all of their non-relational representational features? That isn't clear. There is no descriptive difference between the thinker's beliefs about each Bruce. But is there no representational difference? There is a perspectival difference (Sosa, 2010, pg. 351). And given the connection between perspective and action it's reasonable to count perspective as a representational difference.

Pryor offers an example which bypasses this worry. He imagines a creature, "Flugh", with many eyes, each on independently moving "stalks". The visual experiences of these eyes don't combine into a visual field, and Flugh is not kinesthetically aware of the relative location of his eye stalks in a way that would immediately settle the spatial relations between them.

[O]n one day he has two qualitatively matching experiences of a homogeneous sphere, without any presentation of how the spheres are spatially related to each other. Neither experience even seems to be above, or to the right, of the other. As it turns out Flugh is seeing only a single sphere. In fact, it may be that he's only seeing a single sphere with a single eye, but through some signaling glitch in his brain, he now has cognitively distinct mental presentations." (2016, pg. 331-2)

We have a lack of coordination without any difference in non-relational representational features. And the structure of Flugh's perceptual system makes it difficult to characterize any sense in which the presentations are from different perspectives

3.3 Intransitivity

This thought might tempt Fregeans: Suppose coordination is explained by a relation, R, holding between occurrences. Maybe this doesn't conflict with FC. We could define senses as equivalence classes of occurrences. In particular, we could define the sense of an occurrence o as the equivalence class of occurrences which are R-related to o (see (Heck, 1995), (Rattan, 2009), (Dickie and Rattan, 2010), (Cumming, 2013)). Whether this counts as relationism depends on the details of the intrinsic/relational distinction, and how we characterize representational features.

One Relationist argument undercuts even this form of Fregeanism. Identifying sense with equivalence classes of occurrences presupposes that coordination is an equivalence relation. But some relationists have argued coordination is intransitive, and thus not an equivalence relation. Pinillos offers examples of intransitive coordination. Consider (4) (2011, pg. 314).

(4) We were debating whether to investigate both Hesperus₁ and Phosphorus₂; but when we got evidence of their true identity, we immediately sent probes there_{1/2}

Sharing a subscript indicates coordination. Although Hesperus is coordinated with 'there', and 'Phosphorus' is coordinated with 'there', 'Hesperus' and 'Phosphorus' are not coordinated with each other. Thus the proposed coordination structure is intransitive (for critical discussion, see (Goodsell, 2014), (Contim, 2016)).

To see why this coordination structure is plausible, note that a speaker who believed that 'Hesperus' and 'there' referred to different things would have misunderstood (4). The same goes for 'Phosphorus' and 'there'. But a speaker who doubts that 'Hesperus' and 'Phosphorus' corefer might have understood (4) perfectly. They might be in the grip of an astronomical error. Such a person would think that though 'Phosphorus' and 'Hesperus' each refers, 'there' fails to refer (or, perhaps, fails to have a unique referent).

Paderewski-cases also exhibit intransitivity in the interaction between intra-personal and inter-personal coordination (Taschek, 1998, pg. 336), (Fine, 2009, pg. 119). Peter's friend, Masha, is not confused about Paderewski (she believes there is one man). Peter and Masha agree that Paderewski is musical. So those beliefs are coordinated. They agree that Paderewski is a statesman. So those beliefs are coordinated. But Masha's beliefs are coordinated with each other and Peter's are not. Transitivity of coordination fails (for another example of intransitivity as the product of interpersonal and intrapersonal coordination, see (Richard, 1990, pg. 210)).

3.4 Degrees of Freedom

CSC holds that nothing beyond referential content and coordination determines cognitive significance. But for the Fregean, two bodies of information which have the same referential content, and the same coordination-structure, can still differ in significance. This is because an assignment of senses to a body of referential content determines a coordination-structure, but a coordination-structure doesn't determine an assignment of senses. Take a distribution of senses d which determines a coordination structure r. There will be a distinct distribution of senses d which also instantiates r - simply imagine a permutation of d which maintains the patterns of re-occurrence of sense.

Can any such d and d'differ in cognitive significance? Can there be rationalizing explanations which apply to d but not to d'? Take all of Sally's attitudes about Dylan, and divide them into two coordinated bodies (think of them as the Dylan-body and the Zimmerman-body). Can we imagine a permutation of her mental state which didn't change referential content, didn't change coordination relations, but changed the rationalizing explanations which applied to her? Is there an alternative mental state in which the two bodies of coordinated content remain the same but simply 'switch' senses?

It is plausible that permutations of this kind don't correspond to any cognitive reality (if we confine ourselves to third-personal beliefs, non-demonstrative belief). Fregeans are committed to such permutations describing genuine cognitive differences; relationists aren't (*Fine*, 2009, pg. 59-60), (*Heck*, 2012, pg. 150),(*Pryor*, 2016, pg. 334). For relationists, the Fregean account has too many degrees of freedom.

4 Varieties of Relationism

Relationism holds that occurrences of expressions, or elements of content, are coordinated *iff* they stand in a certain relation. Developing Relationism requires characterizing this relation.

We can distinguish two strategies.

4.1 FORMAL RELATIONISM

Formal Relationism (**FR**) has its roots in *syntactic* solutions to Frege's Puzzle. Beginning with Putnam (1954), theorists - *e.g.*, Fodor (1990), Kaplan (1990), Feingo and May (2006) - have explained the difference in cognitive significance between referentially-equivalent representation in non-semantic terms. **FR** is an elaboration of this idea.

Consider an argument inspired by (*Mates*, 1952). Suppose 'doctor' and 'physician' are synonyms. This entails, given standard assumptions about compositionality, that substituting 'physician' for 'doctor' in any complex expression will not change the significance of the larger expression. It means that (5a) has the same significance as (5b).

- (5) (a) All doctors are doctors
 - (b) All doctors are physicians

Can someone who understands (5a) and (5b) rationally take different attitudes towards them? Take for granted that no one who understands (5a) could rationally reject it. Certainly someone might reject (5b); someone might think that physicians must be family doctors, for example. It's less clear whether this sort of mistake would betray a lack of understanding of 'physician'; it isn't easy to distinguish factual mistakes from mistakes of meaning. Remaining agnostic about that question, we can leverage uncertainty about it to demonstrate a difference in significance between (5a) and (5b) by embedding them in larger constructions.

- (6) (a) Whoever believes that all doctors are doctors, believes that all doctors are doctors
 - (b) Whoever believes that all doctors are doctors, believes that all doctors are physicians

(6 a) and (6 b) should have the same significance. Again, speakers can take different attitudes towards them. But to imagine someone who rejects (6 b) we needn't imagine someone with strange beliefs about the extension of 'physician'. All we need to do is imagine a person with the second-order belief that one can understand 'physician' and 'doctor' without believing that they are coextensive. And surely having the correct belief about that, whatever it is, cannot be a condition on understanding (6 b). So it's hard to avoid the conclusion that (5 a) and (5 b) differ in significance.

We might take this to show that 'doctor' and 'physician' are not synonyms. But on reflection, it's unclear that any pair of distinct words, when substituted for 'doctor' and 'physician' in (6 a) and (6 b), would generate sentences with the same significance. The considerations which made us treat (6 a) and (6 b) differently didn't involve suspicions about the semantic difference between 'doctor' and 'physician'. Rather, it seemed to depend on the difference between repetition of a single word, and the occurrence of distinct words. When a word is repeated, the content of the occurrences is coordinated⁹; when different words are used, the content need not be. To put it a different way, a formal relation holds between the expressions which occur in 'all ... are' in (5 a), which does not hold between the expressions in the same position in (5 b). And the presence of this formal relation results in coordination.

⁹This is probably too strong. On a natural understanding, two occurrences of the same word can be uncoordinated: consider Paderewski cases. One would have to introduce a more abstract conception of expression-type to hold onto the strong generalization (see (*Fiengo and May*, 2006)). At least this much is true: the difference between repetition of a single world and occurrence of different worlds is strongly relevant to coordination, and **FR** takes this as revelatory of the nature of the phenomenon.

The lesson is that the cognitive significance of a representation depends on relational features of its structure (*Putnam*, 1954), (*Taschek*, 1995) (for similar points made in relation to different kinds of cases, see (*Richard*, 1983, 1987)). We can't trace the difference in significance between (5 a) and (5 b) to any semantic difference between 'doctor' and 'physician'. The cognitive difference between (5 a) and (5 b) is a structural feature of the sentences as wholes.

FR generalizes this idea in two ways. First, (6 a) and (6 b) involve formal relations within a particular representation. FR extends this idea to include formal relations that hold across representations (i.e. across sentences or attitudes). If we want to explain the cognitive difference between (1a)["Bob Dylan was born in Minnesota."] and (1b)["Robert Zimmerman was born in Minnesota."] we cannot appeal to any local difference in logical form. Considered by themselves, (1a) and (1b) have the same form. But if we consider them in relation to other representations, we can recognize a 'global' difference in logical form (Taschek, 1995, pg. 82-83). For example, the occurrence of 'Dylan' in (1a) is formally related to the occurrence of 'Dylan' in (7a), but not to the occurrence of 'Zimmerman' in (7b).

- (7) (a) After a motorcycle accident, Dylan moved to upstate New York to make acoustic music.
 - (b) In 1961 Zimmerman moved from Minnesota to NYC.

The second way that **FR** elaborates on the syntactic approach is to insist that the formal features which partially determine cognitive significance are *relational primitives* in rationalizing explanation. To explain this, I'll use intra-personal mental coordination as an example.

On Fodor's (1990) view, belief is a relation between an agent, a proposition, and a sentence in the language of thought. An object occurs in the referential content of a belief-state in virtue of the occurrence of a mental name in the language of thought. Distinct mental names can have the same referent; and thus distinct belief-states can share referential content. Elements of referential content are coordinated *iff* they are introduced by tokens of the same mental name.

Fodor's account is a syntactic approach to Frege's puzzle. But it presupposes his computational theory of mind. FR accepts the spirit of Fodor's account, but resists replacing talk of formal relations with talk of re-occurrence of mental names. That would conflate a question about cognitive relations with questions about the grounding of cognitive relations (Heck, 2012, pg. 153), (Pryor, 2016, pg. 337). Just as there are questions about in virtue of what an agent is in a mental state with a certain referential content - questions which have competing answers - there are questions about in virtue of what an agent is in a state with formal relations holding among elements of referential content (this question, too, admits of competing answers). FR aims to add to our conception of the vocabulary in which rationalizing explanation takes place. In addition to attitude type (belief, desire, etc), and referential content, rationalizing explanation makes direct appeal to formal relations among representational vehicles (Heck, 2012, pg. 154). It might be that in the intra-personal mental domain, formal relations are grounded in the computational structure of mental vehicles. But it's unlikely that explanation could be extended to formal relations in other domains. And we should distinguish the vocabulary involved in a given domain of explanation from substantive theories about the metaphysics of that domain.

We're left with the following:

Formal Relationism The cognitive significance of a representation is determined by its referential content *and* the formal relations which hold between the vehicles of that content and other representational vehicles.

FR carries significant commitments about the relation between content, vehicle, and rationalizing explanation. Though **FR** doesn't commit to any particular view about the non-semantic structure of representational states, it is committed to there being enough non-semantic structure to instantiate an adequate range of formal relations. Language has this much structure in its vehicles but it is not universally accepted that thought does.

Relatedly, **FR** accepts that agents can have distinct belief-states with the same content (*Heck*, 2012), (*Pryor*, 2016). So it must reject the attractive idea that rationalizing explanation is intentional explanation: that it subsumes token propositional attitudes merely in virtue of their mode (belief, desire, etc) and their content (*Fodor*, 1995). For **FR**, distinct belief-states with the same content can play different roles in rationalizing explanation in virtue of standing in different formal relations to other states. This is **FR**'s solution to Frege's puzzle.

We shouldn't understate the importance of this. We might have thought that it was essential to rationalizing explanation that it was *intentional*. Isn't this, in part, what distinguishes it from other sorts of explanation? **FR** must hold that rationalizing explanation is, in a sense, hybrid: appealing both to semantic and non-semantic properties of representations. It must hold that there is no level of explanation which is purely content-driven; all psychological explanation is infused with appeal to the way that content is carried by non-semantic vehicles. For **FR**, Frege's Puzzle provides the materials for a transcendental deduction of the structure of representational vehicles. It is a condition on the possibility of rationalizing explanation that the vehicles with which we represent the world - our sentences and attitudes - have enough structure to instantiate formal relations.

4.2 Semantic Relationism¹¹

Where \mathbf{FR} posits representationally relevant non-semantic relations, Semantic Relationism (\mathbf{SR}) posits *semantic* relations. Thus \mathbf{SR} accepts that cognitive significance is determined by semantic content.

We can introduce **SR** by starting with variables. Fine (2009, Chp 1) claims that our semantics should validate seemingly inconsistent intuitions about variables: that (8 a) and (8 b) have the same meaning; and that (9 a) and (9 b) have different meanings.

- (8) (a) x > 0
 - (b) y > 0
- (9) (a) x > x
 - (b) x > y

(8 a) and (8 b) are notational variants, they do not differ in meaning. It follows - given assumptions about synonymy and compositionality ($Pickel\ and\ Rabern$, 2016) - that 'x' and 'y' have the same meaning. But (9 a) and (9 b) are not synonymous; (9 b) is satisfiable on the intended interpretation, (9 a) is not. Given they differ only in the substitution of an occurrence of 'x' for an occurrence of 'y', it follows that 'x' and 'y' have different meanings.

¹⁰Compare: (Davies, 1992; 1998; 2004), (Rey, 1995).

¹¹I focus on Fine. See notes (13) and (14) for places where other versions of Semantic Relationism diverge. The distinction between Formal and Semantic Relationism is not exhaustive of relationist approaches. There are approaches which accept **CSC** and **RC** but don't fall straightforwardly into either category. Schroeter (2012) argues that meanings should be individuated relationally, but thinks of this as a claim about meta-semantics rather than semantics. It might be that her view is a version of equivalence-class Fregeanism.

Fine solves the antinomy by distinguishing intrinsic semantic properties of variables from semantic relations that hold between occurrences of variables. He holds that 'x' and 'y' have the same semantic properties: they have the same domain. But the pair of occurrences of 'x' in (9 a) stand in a different semantic relation than does the pair 'x', 'y' in (9 b). The pair of occurrences of 'x' in (9 a) must take the same value from the domain on each assignment; 'x' and 'y' in (9 b) can take different values relative to an assignment. Fine argues that the only way to secure this difference between the pairs 'x', 'x' and 'x', 'y', without the unwanted entailment that (8 a) and (8 b) differ in meaning, is to hold that the semantic relation between 'x', 'x' doesn't hold in virtue of the intrinsic semantic properties of each occurrence. That is, we must hold that there are semantic relations which are not reducible to semantic properties. The rules for the language must, in addition to assigning semantic properties to individual occurrences of variables, assign semantic relations to sequences of variables (for critical discussion, see (Pickel and Rabern, 2016, 2017)).

To address Frege's Puzzle, we must extend this account. How can the distinction between intrinsic and relational semantic features illuminate the cognitive difference between coreferential names? Names don't have a variable meaning.¹² The intrinsic semantic feature of a name is its referent; if the referent of a name is intrinsic, how does this leave room for semantic relations to do any work?

To solve this problem, Fine introduces strict coreference.¹³ In addition to the intrinsic semantic requirement that a name-occurrence n_1 refer to an object o, there might be a semantic requirement that name-occurrences n_1 , n_2 refer to the same thing. For Fine, it can be a semantic requirement that n_1 refers to o and a semantic requirement that n_2 refers to o, while failing to be a semantic requirement that n_1 and n_2 corefer (2009, pg. 43). This is the case with respect to 'Dylan' and 'Zimmerman' in (10 a). The occurrences of 'Dylan' in (10 b), on the other hand, are semantically required to corefer.

- (10) (a) Zimmerman is taller than Dylan
 - (b) Dylan is taller than Dylan

Strict coreference doesn't, by itself, show how to extend **SR** to names. We need to know how semantic relations on sequences of names affect the semantic properties of larger expressions. With variables, because the semantic relations between variables determines the values that variables can take relative to assignments, semantic relations affect truth-conditions. But given the semantic *properties* of names determine reference, semantic *relations* between names cannot at least in extensional contexts - affect truth-conditions.

So **SR** posits a non-truth-conditional effect of strict coreference at the level of the sentence or discourse (*Fine*, 2009, pg. 54), (*Pinillos*, 2011, pg. 318). Fine introduces coordination at the level of propositional content. Coordination, in this sense, is the semantic effect of strict coreference: if name-occurrences are strictly coreferential, the elements of content they contribute are coordinated.

It is important not to confuse this notion of coordination with the one we introduced in section (2). That notion is *descriptive*. It describes a class of logical relations which representations

¹²At least they don't on standard accounts of names. But see, e.g., (Cumming, 2008).

¹³Other relationists introduce different term-level semantic relations. Pinillos (2011) introduces 'p-linking', a relational semantic primitive characterized by axioms. Lawlor (2010) notes that there are cases where coordination holds between distinct objects, or between thoughts without objects. She argues that fundamental semantic relation should be consistent with failure of coreference. Fine responds in (2010b).

can stand in. The notion here semantically encodes this descriptive notion. **SR** holds that representations are coordinated (in the descriptive sense) just in case the elements of content they introduce are coordinated (in the semantic sense) at the level of content.

For Fine, a sequence of representations determines a sequence of Russellian propositions: these represent objects bearing properties and standing in relations. A sequence of representations also determines a coordination-relation on elements of its Russellian content. Formally, a coordination-relation is a set of sets of elements of referential content; informally, for each object which occurs anywhere in a sequence of propositions, a coordination-relation determines which occurrences of that object are 'represented as the same'. It's important that a body of representations determines a sequence of propositions rather than a set. We must allow that the same proposition can occur more than once in the body of information, because distinct representations can have the same Russellian content while differing in coordinated content.

The content of the sequence (11) is captured in (12a) and (12b).

- (11) Zimmerman was born in Minnesota. Dylan was born in Minnesota. Dylan is a folk singer
- (12) (a) <<Dylan₁, born in, Minnesota>, <Dylan₂, born in, Minnesota>, <Dylan₃, folk singer>>
 - (b) $\{\{Dylan_1\}, \{Dylan_2, Dylan_3\}\}$

(12 a) is a sequence of Russellian propositions. The same proposition occupies the first two places in the sequence. The subscripts are a metalanguage device for keeping track of Dylan as he appears in difference places in the sequence. (12 b) is a coordination-relation. It captures the fact that the second and third Dylan-elements are coordinated with each other, but neither is coordinated with the first.

Semantic coordination is a representational primitive: different coordination relations have no effect on the truth-conditions of a body of representations. Ultimately, **SR** suggests that Frege's puzzle requires that we introduce a semantically primitive notion of representation as the same.¹⁵

Semantic Relationism Propositional content is irreducibly relational. The cognitive significance of a representation is determined by its referential content *and* semantic relations which hold among elements of that content.

Semantic coordination has dramatic effects on semantic theory. Because coordination relations aren't determined by intrinsic semantic properties, coordinated content is non-compositional in the following respect¹⁶: the coordinated content of a sequence of representations isn't determined by the coordinated content of its subsequences. The coordinated content of s_1 , s_2 , s_3 , s_4 is not determined by the coordinated content of s_1 , s_2 and the coordinated content of s_3 , s_4 , because

¹⁴This way of modeling semantic coordination presupposes coordination is transitive. Fine (2009, Chp 4) accepts this in the intrapersonal case, but denies it in the interpersonal case. Pinillos (2011) rejects transitivity in the intra-personal case, but doesn't offer a model of semantic coordination.

¹⁵It might seem that we've given substance to semantic coordination by claiming that it is the sentence-level effect of strict-coreference between terms. But this is an illusion. First, it gets things the wrong way around: the representational import of a sub-sentential semantic feature is given by how it affects the semantic features of sentences or discourses containing it. Second, Fine wants semantic coordination to play a role in the content of attitudes: a body of coordinated propositions gives the content of an agent's beliefs (2009, Chp 3). But he denies that he is thereby committed to there being non-semantic structure in the attitudes (ibid. pg. 73).

¹⁶For discussions of Relationism and compositionality see (*Taschek*, 1995), (*Fine*, 2009, pg. 26), (*Pickel and Rabern*, 2017).

the coordinated content of s_1 , s_2 , s_3 , s_4 depends on relations that can hold between expressions in s_1 and s_4 .

For \mathbf{SR} , the basic unit of interpretation is a sequence of sentences. A form of compositionality holds at the level of sequences: given a sequence $s_1, ..., s_n$, substituting an expression e for an expression e' which agrees with e in its semantic relations to every expression in $s_1, ..., s_n$ will not alter the content of $s_1, ..., s_n$. But \mathbf{SR} enforces a form of top-down interpretation which is unorthodox. Interpreting a sequence of sentences requires re-interpreting the entire sequence at each stage, because the coordinated content of the sequence is not determined by the coordinated content of an initial segment and the coordinated content of the later segment (Fine, 2009, pg. 83-84).

Treating coordination as a semantic phenomenon opens up possibilities for the meaning of hyper-intensional operators. **SR** holds that the truth of 'S believes that P' depends not only on the referential content of 'P' but on its coordinated content (Fine, 2009, chp 4). This is straightforward when we consider relations internal to 'P'. We can explain the fact that the truth-conditions of (13 a) and (13 b) differ by appealing to the fact that 'Dylan is taller than Zimmerman' and 'Dylan is taller than Dylan' have different coordinated content.

- (13) (a) Sally believes that Dylan is taller than Zimmerman
 - (b) Sally believes that Dylan is taller than Dylan

But it's less clear what to say about differences that appeal to external relations. (14 a) and (14 b) can differ in truth-value.

- (14) (a) Sally believes that Dylan is a poet
 - (b) Sally believes that Zimmerman is a poet

But the contained sentences do not differ in internal relations; they only differ in content when considered as elements of larger sequences (*Soames*, 2010). Semantic Relationists have tried different ways to turn these external differences into truth-conditional differences at the level of attributions (see Fine (2009, chp. 4), (2010a) and Pinillos (2015)). It isn't clear that any plausible compositional semantics is possible.¹⁷

It isn't an accidental feature of **SR** that it offers no straightforward way of giving a compositional semantics for attitude ascriptions. Just as **SR** rejects an atomistic conception of meaning, it rejects an atomistic conception of the attitudes. **SR** treats belief as a relation to a coordinated proposition. Coordinated propositions are individuated in part by their links to other propositions (cf. Fine's (2010a) notion of the 'tertiary' level of content). An individual's total belief-state isn't an aggregate of her relations to individual propositions. Given this, belief-attributions will be interpreted relative to an implicit background of coordinated propositions which are not the semantic value of anything in the attribution.

5 LOOKING AHEAD

Relationism promises a new approach to Frege's puzzle. Both Formal and Semantic varieties of Relationism posit new representational primitives: either in the relations between vehicles or in the structure of content. Given this, it is not easy to see how we should chose between them. It

¹⁷For an anticipation of Relationist approaches to attitude ascriptions, see (*Richard*, 1990, pp. 173ff). For an account of the semantics of attitude ascriptions in the spirit of FR, see (*Cumming*, 2013).

seems likely that progress will come from exploring the different way that the two forms of Relationism depart from traditional assumptions about content, compositionality, representational vehicle, and rationalizing explanation.

Whether or not we decide to accept some version of Relationism, Relationist approaches clarify the structure of Frege's puzzle and the space of possible solutions. Going forward, Relationism should be considered one of the options for solving Frege's puzzle.

References

Almotahari, M. (2013), Flaws of formal relationism, Thought: A Journal of Philosophy, 2(4), 367–376.

Burge, T. (1979), Sinning against Frege, The Philosophical Review, 88(3), 398-432.

Campbell, J. (1987), Is sense transparent?, in *Proceedings of the Aristotelian Society*, vol. 88, pp. 273–292, Aristotelian Society.

Contim, F. D. V. (2016), Mental files and non-transitive de jure coreference, Review of Philosophy and Psychology, 7(2), 365–388.

Cumming, S. (2008), Variabilism, Philosophical Review, 18, 525–554.

Cumming, S. (2013), From coordination to content, *Philosophers' Imprint*, 13(4), 1–16.

Davies, M. (1992), Aunty's own argument for the language of thought, in *Cognition*, *Semantics and Philosophy*, pp. 235–271, Springer.

Davies, M. (1998), Language, thought and the language of thought (Aunty's own argument revisited), in Language and thought: interdisciplinary themes, edited by P. Carruthers and J. Boucher, pp. 226–247, Cambridge University Press.

Davies, M. (2004), Aunty's argument and armchair knowledge, in *Language*, *Knowledge*, and *Representation*, pp. 19–37, Springer.

Dickie, I., and G. Rattan (2010), Sense, communication, and rational engagement, dialectica, 64(2), 131–151.

Dummett, M. (1981), Frege: Philosophy of Language, Harvard University Press.

Evans, G. (1982), The Varieties of Reference, Clarendon Press.

Evans, G. (1985), Understanding Demonstratives, in Collected Papers, pp. 291–321, Clarendon.

Fiengo, R., and R. May (2006), De lingua belief, The MIT Press.

Fine, K. (2009), Semantic Relationism, Wiley-Blackwell.

Fine, K. (2010a), Comments on Scott Soames "Coordination problems", *Philosophy and Phenomenological Research*, 81(2), 475–484.

Fine, K. (2010b), Reply to Lawlor's "Varieties of Coreference", *Philosophy and Phenomenological Research*, 81(2), 496–501.

Fodor, J. (1990), Substitution arguments and the individuation of beliefs, in *Meaning and method: Essays* in honor of Hilary Putnam, edited by G. Boolos, pp. 63–79, Cambridge University Press.

Fodor, J. A. (1995), The elm and the expert: Mentalese and its semantics, MIT press.

Frege, G. (1892), On sinn and bedeutung, in *The Frege Reader*, pp. 151–171, Blackwell.

Frege, G. (1918), The Thought: A logical inquiry, in *The Freqe Reader*, Blackwell.

Goodsell, T. (2014), Is de jure coreference non-transitive?, Philosophical Studies, 167(2), 291–312.

Grav. A. (2016). Minimal descriptivism, Review of Philosophy and Psychology, 7(2), 343–364.

Heck, R. G. (2014), In defense of formal relationism, Thought: A Journal of Philosophy, 3(3), 243-250.

Heck, R. G., Jr. (1995), The sense of communication, Mind, 104, 79–106.

Heck, R. G., Jr. (2012), Solving Frege's puzzle, The Journal of Philosophy, 109, 132–174.

Heck, R. G., Jr. (ms.), Solving Frege's Puzzle (extended version), http://rgheck.frege.org.

Kaplan, D. (1990), Words, Proceedings of the Aristotelian Society.

Kremer, M. (2010), Sense and reference: The origins and development of the distinction, in *The Cambridge Companion to Frege*, edited by T. Ricketts and M. D. Potter, pp. 220–292, Cambridge University Press.

Kripke, S. (1979), A puzzle about belief, in *Meaning and Use*, edited by A. Margalit, pp. 239–283, D Reidal.

Kripke, S. (1980), Naming and Necessity, Harvard University Press.

Lawlor, K. (2010), Varieties of coreference, Philosophy and Phenomenological Research, 81(2), 485–495.

Mates, B. (1952), Synonymity, in Semantics and the Philosophy of Language, edited by L. Linsky, University of Illinois Press.

McDowell, J. (1977), On the sense and reference of a proper name, Mind, 86(342), 159-185.

- Millikan, R. G. (1997), Images of identity: In search of modes of presentation, Mind, 106 (423), 499–519. Peacocke, C. (1983), Sense and content: Experience, thought and their relation, Oxford: Oxford University Press.
- Peacocke, C. (1992), A study of concepts., The MIT Press.
- Perry, J. (1977), Frege on demonstratives, The philosophical review, 86(4), 474-497.
- Pickel, B., and B. Rabern (2016), The antinomy of the variable: A tarskian resolution, *The Journal of Philosophy*, 113(3), 137–170.
- Pickel, B., and B. Rabern (2017), Does semantic relationism solve frege's puzzle?, *Journal of Philosophical Logic*, 46(1), 97–118.
- Pinillos, N. n. (2011), Coreference and meaning, *Philosophical Studies*, 154(2), 301–324.
- Pinillos, N. n. (2015), Millianism, relationism, and attitude ascriptions, in *On Reference*, edited by A. Bianchi, pp. 322–334, Oxford.
- Pryor, J. (2016), Mental graphs, Review of Philosophy and Psychology, 7(2), 309–341.
- Putnam, H. (1954), Synonymity, and the analysis of belief sentences, Analysis, 14(5), 114–122.
- Rattan, G. (2009), Semantic relationism, by Kit Fine., Mind, 118 (472), 1124–1131.
- Rey, G. (1995), A not "merely empirical" argument for a language of thought, *Philosophical Perspectives*, 9, 201–222.
- Richard, M. (1983), Direct reference and ascriptions of belief, Journal of Philosophical Logic, 12(4), 425–452.
- Richard, M. (1987), Quantification and leibniz's law, The Philosophical Review, 96(4), 555–578.
- Richard, M. (1990), Propositional attitudes: An essay on thoughts and how we ascribe them, Cambridge University Press.
- Schroeter, L. (2012), Bootstrapping our way to samesaying, Synthese, 189(1), 177–197.
- Soames, S. (2010), Coordination problems, Philosophy and Phenomenological Research, 2, 464–474.
- Sosa, D. (2010), The Fine line, Analysis Reviews, 00, 1–12.
- Taschek, W. W. (1992), Frege's puzzle, sense, and information content, Mind, 101 (404), 767–791.
- Taschek, W. W. (1995), Belief, substitution, and logical structure, Noûs, 29(1), 71–95.
- Taschek, W. W. (1998), On ascribing beliefs: Content in context, The Journal of philosophy, 95(7), 323-353.
- Taylor, K. A. (2003), What's in a Name?, in Reference and the Rational Mind, pp. 1–32, CSLI Publications.
- Taylor, K. A. (2010), On Singularity, in *New essays on singular thought*, edited by R. Jeshion, pp. 78–102, Oxford Univ Press.