Fodor on Concepts and Frege Puzzles

MURAT AYDEDE
May 1998

The University of Chicago
Department of Philosophy
1010 E. 59th Street
Chicago, IL 60637, USA
`m-aydede@uchicago.edu`

ABSTRACT. Fodor characterizes concepts as consisting of two dimensions: one is *content*, which is purely denotational/broad, the other the Mentalese *vehicle* bearing that content, which Fodor calls the Mode of Presentation (MOP), understood "syntactically." I argue that, so understood, concepts are not interpersonally sharable; so Fodor's own account violates what he calls the Publicity Constraint in his (1998) book. Furthermore, I argue that Fodor's non-semantic, or "syntactic," solution to Frege cases succumbs to the problem of providing interpersonally applicable functional roles for MOPs. This is a serious problem because Fodor himself has argued extensively that if Fregean senses or meanings are understood as functional/conceptual roles, then they can't be public, since, according to Fodor, there are no interpersonally applicable functional roles in the relevant senses. I elaborate on these relevant senses in the paper.

In Chapter 2 of his new book, *Concepts: Where Cognitive Science Went Wrong* (1998), Fodor enumerates five constraints that any theory of concepts, he says, must satisfy. He takes these constraints to be empirical but non-negotiable. Among them is the Publicity Constraint (PC), which says "concepts are *public*; they're the sorts of things that lots of people can, and do, *share*" (1998: 28). In other words,

> [s]ince, according to RTM, concepts are symbols, they are presumed to satisfy a type/token relation; to say that two people share a concept (i.e. that they have literally the same concept) is thus to say that they have tokens of literally the same concept type. The present requirement is that the conditions for typing concept tokens must not be so stringent as to assign practically every concept token to a different type from practically any other." (1998: 28)

Fodor also distinguishes between *concept* identity and *content* identity, where content is understood to be broad (purely denotational). So different contents imply different concepts. But it is possible for two concepts to be type-distinct while identical in content (*intra* or *inter*personally). The extra individuating element is what Fodor calls Modes of Presentation (MOPs), which are the vehicles that carry the content.

So, according to Fodor, the individuation condition for concepts is given by an ordered pair (a 2-tuple), whose first element is the broad content and the second a vehicle type that has the first as its semantic value: <denotation, vehicle type>. For present purposes we can represent concepts with these 2-tuples. According to Fodor, vehicles are terms in one's Language of Thought (LOT — sometimes called Mentalese) realized in the brain. As such vehicles have both syntactic and semantic properties, and this fact can be used to answer the question raised by standard Frege cases: what makes co-denoting concepts type-distinct? In fact, Fodor takes this feature of RTM (= Language of Thought Hypothesis) to provide the theory with a strong empirical support: "The Frege programme needs something that is both in the head and of the right [causal] kind to distinguish coreferential concepts, and the Mates cases suggest that whatever is able to distinguish coreferential concepts is apt for syntactic individuation. Put all this together and it does rather suggest that modes of presentations are syntactically structured mental particulars" (1998: 39, see also his 1989).

In the same book, Fodor also defends an atomistic account of concepts according to which most (lexical) concepts have no internal structures and their content is determined exclusively by their "causal-*cum*-nomological" relations with the world. We may say that the vehicle #dog# has no semantically significant internal (syntactic) structure and expresses the property of *being a dog* (or, just denotes dogs) in virtue of a nomological relation (NR) to dogs. NR then is Fodor's way of naturalizing what it is for the second element in the 2-tuple to have the first as its semantic value. So, for instance, the concept DOG = <*dog*, #dog#>, where the second element has the first as its semantic value in virtue of its standing in NR to it, and is syntactically primitive.

The question I want to raise is whether this atomistic account of concepts Fodor defends against all its non-atomistic rivals itself satisfies the Publicity Constraint. Fodor obviously thinks that it does, because one of his reasons why inferential role accounts of concepts ought to be abandoned in favor of his own account is that they fail to satisfy PC. Over the years Fodor has argued extensively — and in my opinion convincingly — that inferential role semantics (IRS) and other species of functional role semantics (FRS) have holistic consequences which are destructive of intentional psychology. The reason for this, very briefly, is that concepts individuated by their functional roles cannot be shared interpersonally, so they violate PC. I won't challenge this claim.

Now, the question is whether Fodor himself needs, at some point, to individuate concepts in functional terms. This issue arises because, according to Fodor, the naturalistic determination of semantic values of vehicles, i.e. the nature of NR, is

atomistic in the sense that vehicles enter into NR with their semantic values *individually* on the basis of a certain lawlike relation since vehicles are *syntactic atoms*. Assuming that individuation of semantic values themselves is non-problematic, what determines the individuation of syntactic atoms, the vehicles? Fodor needs a non-semantic individuation of vehicles, as his proposed explanation of Frege cases demands. So, on what basis does Fodor propose to type vehicle tokens such that they can be public, i.e. interpersonally sharable? Fodor seems to be baffled by Frege's own solution to Frege cases, and in a way, invites us to share his reaction. For he writes:

> Frege's structural problem is that, though he wants to be an *externalist* about MOPs, the architecture of his theory won't let him. Frege's reason for wanting to be an externalist about MOPs is that he thinks, quite wrongly, that if MOPs are mental then concepts won't turn out to be public. But if MOPs *aren't* mental, what kind of thing *could* they be such that *necessarily* for each MOP there is only one way in which a mind can entertain it! ... If, however, MOPs are in the head, then they can be proximal mental causes and are, to that extent, apt for functional individuation. If MOPs are both in the head and functionally individuated, *then a MOP's identity can be constituted by what happens when you entertain it.* ... Even Frege should have been a mentalist about MOPs if he wished to remain in other respects a Fregean. On the other hand ..., to claim that MOPs must be *mental* objects is quite compatible with also claiming that they are *abstract* objects, and that abstract objects are *not* mental. The apparent tension is reconciled by taking MOPs-qua-things-in-the-head to be tokens of which MOPs-qua-abstract-objects are the types. It seems that Frege thought that if meanings can be shared it somehow follows that they can't also be particulars. But it beats me why he thought so. You might as well argue from '*being a vertebrate* is a universal' to 'spines aren't things'. (1998: 20–1)

Let's just grant the cogency of Fodor's reasoning regarding Frege's own solution. If so, what needs to be provided is a non-semantic method for typing MOP-tokens across different heads. As the passage indicates,[1] Fodor seems to opt for a functional individuation of vehicle tokens for this purpose. In other words, the interpersonal type identity of the second element in 2-tuples is determined functionally. But if this is really his intention, it is equally baffling! For how could *he* be thinking that vehicle tokens can be functionally typed across different heads given that it was actually the unavailability of this method that had led him to conclude that there was no non-holistic type-individuation of functional/computational roles. It was precisely this consequence that made concepts/contents not public on a FRS, according to Fodor, as indeed he continues to argue in the Appendix 5B of the same book (1998). If, as Fodor believes,

---

[1] There are other passages (1998: 19, 22). See also his (1994: 49–50, 105–10).

there are no robust interpersonally sharable functional/computational roles, then there is no non-semantic interpersonal type-individuation of vehicle tokens on the basis of vehicles' functional roles. And, if this is right, then — given the very plausible (and to my knowledge uncontested) claim that physicalism in the form of type-type identity theory is false with respect to *particular* concepts like the concept DOG or propositional attitudes with *particular content* (like the *belief that snow is white*) — there is simply not much left for Fodor to go on in typing Mentalese tokens for purposes of explaining Frege cases.

One seemingly natural reply here might be that Fodor's argument from functionalism to holism (hence non-publicity) depends entirely on functional roles that are taken to be meaning roles (i.e., functional roles relating concepts that are presumed to be meaningfully connected to each other). The slide into holism only affects meaning roles, because it relies on a premise about the failure of the analytic/synthetic distinction. But, one might say, the functional roles used in typing Mentalese tokens in explaining Frege puzzles are not meaning roles, but purely syntactic or orthographic roles. It may be that there is something like a Mentalese alphabet in each head with a common orthography, and different vehicles involved in different concepts correspond to alphabetically distinct types on the basis of their syntactic/orthographic roles, regardless of whether those types have the same meaning roles. Although this reply has a prima facie plausibility, it can't be made to work for an account of interpersonally applicable functional roles. For one thing, the notion of a Mentalese alphabet with an orthography common across heads is totally obscure. The ordinary notion of orthography is more or less given on the basis of letters' physical shapes and forms. If the notion of orthography involved in Mentalese were the ordinary one, this would commit the defender of this reply to a strong from of type-type physicalism of Mentalese vehicles across heads. I take it that this consequence is very implausible and must be rejected (Fodor rejects it: 1994: 105–9). But if the notion of orthography is not the ordinary/formal one, it is not clear what kind of functional specification is appropriate to specify the orthography of, say, #dog# across heads. The only plausible avenue is to take the functional roles of such particular vehicle tokens as specified in terms of their causal relations to other tokens and perhaps to certain perceptual input and behavioral output (all specified non-semantically). These causal roles are to be specified in terms of causal generalizations such as "if S has a B-state that ... #dog#..., then, ceteris paribus, S will tend to have a B-state that ... #animal# ...". But this would amount to having to give an account of such roles in terms of vehicles' functional roles paralleling their "inferential" roles. (In fact this is exactly what Stich (1983) does in the elaboration of his celebrated Syntactic Theory of Mind — for a critical discussion of which see my (forthcoming).) If so, all the problems pertaining to the specification of

such semantic roles can be brought to bear on the interpersonal specification of causal roles of Mentalese tokens. So, I don't think that this "natural" reply works.[2]

But the problem seems more serious: Fodor's own account of concepts seems to require a non-semantic (and non-physical) individuation of vehicle tokens, and it is not clear what this could be within his framework. It is interesting to observe that, on the formulation I have given, Fodorian *concepts* are more or less what a two-factor inferential role semanticist would say *contents/meanings* are: <external denotation, MOPs qua internal functional roles>. To repeat, to prevent misunderstanding: for such a semanticist the tuples represent *contents* of concepts, not *concepts* per se. But this hardly seems to matter. If Fodor opts for functional typing of vehicle tokens across heads, as he seems to be doing, then, contrary to his advertisement, his own account of concepts fails to satisfy PC. And it is difficult to imagine what else he has at his disposal.

The moral, as always, is that Frege cases are likely to cause serious trouble for any one who defends a purely denotational account of mental content. For, as we have seen, going "syntactic" doesn't work even if Fodor is right that going semantic (à la Frege) doesn't work either.

On the other hand, it may be that in the first two chapters of his (1998) Fodor is not being particularly careful about the way he goes back and forth between 'content' and 'concepts'.[3] But it may be argued that what he really means, as can be seen when the text is appropriately re-written, is that concepts are sharable only with respect to their content dimension, i.e. the first element in 2-tuples. So he can consistently maintain that concepts are, *strictly speaking*, not public, but since psychological explanations are given in terms of broad content properties (1994: Chaps. 1–2), making concepts not interpersonally sharable does not pose any threat to a scientific intentional psychology. In other words, to the extent to which vehicles of contents exhibit variations in different heads, to that extent concepts will exhibit variations, but this hardly matters as long as intentional psychological explanations are all broad as Fodor maintains.

One immediate problem with this move is that concepts, strictly speaking, turn out not to be the kind of things we attribute to people in the explanation and prediction of their behavior, including verbal behavior, which is anathema to contemporary cognitive psychology, as indeed seems to be suggested by Fodor himself by the subtitle of his book. One reason why cognitive science went wrong, according to Fodor, is that it treated concepts as non-atomic. But part of the reason why this is wrong is that it

---

[2] For an elaboration of why typing Mentalese tokens non-semantically across heads has exactly the same sort of problems that plague functional role semanticists, see my 1997 and 1998a
[3] Indeed, if you read pp. 28–30 where he explains what PC is and why it is needed, he probably means content when he uses 'concept.' It is very instructive to read these pages while paying attention to his usage of these terms.

tends to make concepts not sharable (especially with Theory-theories of concepts, which are species of FRS). As we have seen, however, Fodor's own theory has the same exact consequence. But put this aside, maybe Fodor really means contents, when he speaks of concepts in relevant contexts.

But a more serious problem with this move is that it makes the explanation of Frege cases non-semantic. In other words, as Fodor himself puts it, "of course, the price I have to pay for this sort of treatment is that you don't get content explanations of cases where the character of a creature's behavior depends on specificities of its MOP" (personal communication). Frege cases are to be treated by adverting to non-semantic differences in the MOPs, i.e. to differences at the level where intentional generalizations are *implemented*, according to Fodor. But there are at least two obvious problems with this. As I indicated elsewhere (1997), when the folk explain why Oedipus married Mother in the usual way, they don't seem to be doing any kind of implementational psychology. On the face of it, the explanation you get — when you hear "well, Oedipus didn't of course know (believe) that Jocasta was Mother" — has nothing to do with adverting to the non-semantic differences in the internal vehicles realized in Oedipus' brain. But secondly, even if we grant that in some sense this is in fact what the folk are doing,[4] there is still the following difficulty for Fodor.

On the face of it, the folk seem not to have any special problem in explaining *intra* as well as *inter*personal Frege cases. For instance, the explanation of why these people ran in the direction of *X* when they were threatened by the perceived danger of *Y* is because they thought that Superman was in that direction to help them. If they only believed Clark Kent was there, they wouldn't run in that direction. These cases are in fact ubiquitous; so such examples can easily be multiplied. If Fodor is right about the Frege cases, it is a mystery how the Folk could be so at ease and successful in their explanation. Certainly, there seem to be robust generalizations involving interpersonal Frege cases. For instance, people feel safer when they believe that Superman is present and act accordingly. State this generalization with 'Clark Kent,' it becomes false. How do the folk manage this if Fodor is right? The puzzle is that if such generalizations make essential reference to people's vehicles, then on Fodor's framework, this should be a mystery, since there is no method of typing vehicle tokens across people: a broad semantic account is out in Frege cases, but so are the physical and functional accounts on Fodor's view. This is the real difficulty apart from what a strict and literal reading of his (1998) suggests.

In fact, this conclusion is more or less accepted by Fodor himself in his (1994) in which he argues that Frege cases like the one involving Oedipus' ignorance of the identity of Jocasta with his mother ought to be taken as exceptions to broad intentional generalizations rather than disconfirming counterexamples to them. (For instance, since

---

[4] For arguments close to this effect, see Crimmins (1992) and Richard (1990).

Oedipus tried to bring about his marriage to his Mom/Jocasta despite his long-standing desire to avoid incest, his case seems to be a counterexample to the psychological generalization that for any S if S desires that P and believes that S can bring about that P, then, ceteris paribus, S will tend to bring about that P.) This is an open acknowledgement that there can be no nomological intentional explanation of Frege cases because psychological generalizations are stated solely in terms of broad content. But since there is no other sort of nomological psychological explanations, i.e. since there is no explanation of cases by subsuming them under law-like generalizations that are applicable interpersonally, it so turns out that, on Fodor's view, there is no nomological explanation of interpersonal Frege cases like the one I have described above, which makes a mystery of how the folk generalize over the Frege cases.[5]

---

[5] For more elaboration on this theme and a discussion of interpersonal Frege cases, see my (1998a) and (1998b). I would like to thank Michael Devitt, Guy Dove, Jerry Fodor, Eric Margolis, Jesse Prinz, Philip Robbins, and Ken Taylor for their helpful comments and suggestions.

REFERENCES:

Aydede, Murat (1997). "Has Fodor Really Changed his Mind on Narrow Content?", *Mind and Language*, Vol. 12, Nos. 3/4, pp. 422–58.

Aydede, Murat (1998a). "Typing Mentalese Tokens," draft, available at
`<http://humanities.uchicago.edu/faculty/aydede/Typing.html>`

Aydede, Murat (1998b). "Language of Thought Hypothesis: State of the Art," draft, available at
`<http://humanities.uchicago.edu/faculty/aydede/LOTH.SEP.html>`
(A shorter version has appeared in *Stanford Encyclopedia of Philosophy*, edited by Edward Zalta, CSLI Publications, Stanford: `<http://plato.stanford.edu/>`)

Aydede, Murat (forthcoming). "Computation and Functionalism: Can Psychology Be Done 'Syntactically'?" in *Synthese*.

Crimmins, Mark (1992). *Talk about Beliefs*, Cambridge MA: MIT Press.

Fodor, Jerry A. (1989). "Substitution Arguments and the Individuation of Belief" in *A Theory of Content and Other Essays*, J. A. Fodor, Cambridge, MA: MIT Press, 1990. (Originally appeared in *Method, Reason and Language*, G. Boolos (ed.), Cambridge, UK: Cambridge University Press, 1989.)

Fodor, Jerry A. (1994). *The Elm and the Expert: Mentalese and Its Semantics*, Cambridge, MA: MIT Press.

Fodor, Jerry A. (1998). *Concepts: Where Cognitive Science Went Wrong*, Oxford, UK: Oxford University Press.

Richard, Mark (1990). *Propositional Attitudes*, Cambridge, UK: Cambridge University Press.