

18. Propositional attitudes and the language of thought

Martín Abreu Zavaleta

June 24, 2014

1 What a good theory of propositional attitudes should look like

Fodor presents a set of requirements on a good theory of propositional attitudes, and offers some defense of them. These are the requirements:

- (I) Propositional attitudes should be analyzed as relations. Attitude verbs express relations between a subject of the attitude and something else. Part of the reason to accept this view comes from apparent syntax: Fodor is assuming that a sentence like ‘John believes that Mary is coming’ has the same syntactic structure as ‘Mary ate my lettuce’, which has a noun phrase in its top node, and a predicate formed of a verb and a noun phrase [NP₁[V NP₂]]. Since these kind of sentences are usually taken to express relations between the two noun phrases, Fodor takes this to be good evidence that sentences ascribing propositional attitudes should also express relations. He offers some further evidence supporting his claim:
 - a. Unlike other superficially relational expressions, it’s not clear how we could translate sentences ascribing propositional attitudes into sentences that are not committed to the existence of a denotation for the that-clause.
 - b. In general, existential generalization applies to the syntactic objects of verbs of propositional attitudes. For instance, if it is true that A believes that dogs are filthy, it follows that there is something that A believes.¹
 - c. The relational view of attitude ascriptions is the only game in town. The alternative, as far as Fodor knows, is the view that sentences like ‘A believes that p’ ought to be syntactically parsed as ‘A believes-that-p’, where the predicate here has no further structure. Here are some considerations against such a view:
 - i. The productivity of belief attributions.
 - ii. Different propositional attitudes can have the same content. E.g. someone can both fear and believe that there’s a monster in the closet.
 - iii. There can be logical relations between beliefs.
 - iv. There is a correlation between indicative sentences of English and the things that seem to be embedded in that-clauses. The fusion view doesn’t explain this correlation.

¹Existential generalization is the inference from a claim about a specific individual to a claim about the existence of a thing with the property predicated of that individual. For instance, from ‘A is F’, we can infer ‘There is at least one thing that is F’.

- v. The relation between the truth of a belief and the truth of the sentences we use to characterize that belief.
- (II) A theory of propositional attitudes should explain the parallelism between verbs of propositional attitudes and verbs of saying. The things that can be said to be believed are the same kinds of things that can be said to be said. Fodor states this condition as arising from syntax, but we can make a distinction between two different conditions, one linguistic and one corresponding to the content of beliefs themselves. The linguistic constraint comes from the observation that the syntactic behavior of ‘said that’ and ‘believes that’ is the same. The observation about content comes from the fact that, in some cases, the following sentences are true: ‘I believe what she said’, ‘I desire the same thing that she fears’, and the like. That these sentences are true seems to imply that the content of sayings is the same kind of thing as the content of propositional attitudes.
 - (III) A theory of propositional attitudes should account for their opacity. We can’t substitute expressions with the same referent in belief attributions *salva veritate*. Again, we can distinguish between a linguistic constraint and a constraint related to the content of beliefs: the linguistic constraint requires us to figure out a semantics for attitude ascriptions that allows for the failures of substitutivity. The constraint on content requires us to explain what the content of propositional attitudes is, such that we can describe it by means of those sentences.
 - (IV) “A theory of propositional attitudes should legitimize the ascription of form to objects of propositional attitudes. In particular, it should explain why the form of a belief is identical to the logical form of the correspondent of a sentence which (opaquely) ascribes that belief.” (p. 546) Fodor notes that the contents of propositional attitudes must be non-arbitrarily related to the causal chains that explain our behavior. This is crucial for stating folk-psychological laws of the following kind: if x believes that A is an action x can perform, and if x believes that the performance of A is sufficient to bring it about that Q, and if x wants it to be the case that Q, then x acts in a fashion intended to be a performance of A. According to Fodor, these generalizations are specified by reference to the form of the content of a propositional attitude. So, if propositional attitudes are to play any role in the explanation of our behavior, they better have the same form as the kind of statements that we use to state such laws. He adds two further reasons to accept this condition:
 - a. We may want to compare beliefs with respect to their form: A’s belief that p and q is the conjunction of B’s belief that p and C’s belief that q.
 - b. We may wish to speak of the logical consequences of a belief, even though we admit that the consequences of a belief may not be themselves the objects of a belief.
 - (V) A theory of propositional attitudes should mesh with empirical accounts of mental processes.

2 Carnap’s theory

Fodor observes that Carnap’s theory does surprisingly well in satisfying the criteria offered above. Let’s start by stating Carnap’s theory:

Carnap's theory: propositional attitudes are relations between people and the sentences they are disposed to utter, e.g. between people and sentences of English. In particular, to believe that P is to be disposed to utter tokens of the sentence 'P'.

Let's go one by one:

- I. It's easy to see that, on Carnap's view, propositional attitudes are analyzed as relations between the subject of the attitude and something else: sentences. It's also easy to see how if a belief is true, then so is the sentence embedded under the that-clause, since the sentence embedded is what is believed.
- II. On this account, saying that P and believing that P are relations to tokens of the same sentence, so there is no mystery as to how these two kinds of sentences are syntactically similar.
- III. On this account, attitudes are relationships to sentences syntactically individuated, not to propositions or the like. So it's not surprising that we can't substitute one sentence for other with the same truth value while preserving truth: someone may bear the required relation to one sentence, but not to the other.
- IV. On Carnap's account, the logical form of the object of a propositional attitude is just the logical form of the sentence that we embed in the belief report.
- V. An example of the kind of empirical explanations that Carnap's view can explain is the following. We can start with the observation that some beliefs are more complex than others, and that the complexity of a belief seems to be correlated with the complexity of the sentences we use to describe it. We would have an explanation of this fact if our account of the content of propositional attitudes allowed us to predict the complexity of a mental state from the complexity of whatever the theory assigns as the object of the attitude. Carnap's theory can easily explain this fact, since it can claim that the complexity of a mental state depends on the complexity of the sentence that constitutes the object of the belief relation. This in turn is related to the complexity of the sentences we use to describe propositional attitudes.

Thus, Carnap's theory is pretty good at satisfying the requirements, but it's not perfect. Fodor points out the following problems:

- (i) Carnap's view is behavioristic. But there are plenty of reasons to reject behavioristic accounts of propositional attitudes. Notice that this only affects Carnap's account of the relation between subjects and sentences, but it doesn't require modifying the view about the objects of the relation. For instance, Carnap could reformulate the view along functionalist lines while keeping the claim that the objects of the relation are token sentences of English.
- (ii) Carnap's view is too fine-grained. Plausibly, 'Sarah believes that Suárez bit Chiellini' and 'Sarah believes that Chiellini was bitten by Suárez' ascribe the same belief. However, the sentences embedded under the that clauses of these two ascriptions are different, so Carnap's view predicts that the beliefs they ascribed are different too.
- (iii) Someone can believe that it's raining even though she doesn't speak English.
- (iv) Someone can believe that it's raining even if she doesn't speak any language at all.

- (v) Some beliefs may be inexpressible. Carnap's view can't make room for this. What could such a belief be? Suppose that space is Newtonian (there is an infinite amount of space and there is no privileged point in it). Then, there is a sense in which you don't know where you are in all that space. But, as it happens, this kind of ignorance can't be expressed (see Dasgupta 2014) by any sentence in English (or at least it's not clear how one would express it). Carnap is committed to thinking that it's not true that one doesn't know his position in Newtonian space. Needless to say, this should not be considered a devastating argument.
- (vi) If token sentences are the objects of propositional attitudes, how are first languages learned? Presumably, the process of learning a language involves generalizations on the data, formulation of hypotheses, etc. All of these must involve beliefs, expectation, and the like. So, on pain of circularity, there must be some propositional attitudes which are not relations to sentences of English.

3 The Language of Thought to the rescue

Most of the problems with Carnap's view come from its use of sentences of natural language. Fodor proposes to correct the theory by taking the objects of propositional attitudes to be sentences (formulas) of a non-natural language, or an Internal Representational System (a language of thought). He then argues that this sort of view can satisfy the conditions he has offered so far.

Condition I is satisfied merely by the fact that propositional attitudes are analyzed as relations. Condition III is satisfied in pretty much the same way as Carnap's view satisfied it, with the sole difference that the sentences in question now are sentences in a language of thought.

Conditions II and IV can be seen as imposing the requirement that the sentences in the language of thought that we ascribe relations to have the same truth conditions as the sentences in natural language that we use to attribute those relations. The *that*-clause in the attitude ascription connects the ascription to a formula in the internal language of the believer, but it also connects that internal formula with a sentence in natural language: the sentence embedded in the ascription by means of the *that*-clause.

The idea is that the conventions of natural language establish pairings of natural language sentences with sentences in a language of thought. This is compatible with the individuation of a given formula in the language of thought by means of the functional role that it has in the production of utterances of the sentence it is related to.

Fodor claims that this view escapes the objections presented against Carnap's view on the following assumption:

Let's assume that the internal language is innate, that its formulae correspond one-one with the contents of propositional attitudes, and that it is *as* universal as human psychology; viz. that to the extent that an organism shares our mental processes, it also shares our system of internal representations. (p. 551)

On this assumption, the objections above don't apply to the language of thought hypothesis. **Question:** Why?

Still, some people may think that this hypothesis is just implausible. In its defense, Fodor claims that our best psychological theories are committed to the truth of the hypothesis anyway. He illustrates as follows. Consider the ambiguous sentence 'they are flying planes'. The ambiguity of this

sentence is usually understood to be derived from the fact that there are two ways of understanding the syntactic structure of the sentence. This in turn is usually understood as there being a function from strings of symbols to syntactic structures. But what, asks Fodor, determines that some such function is the correct one? It's not very explanatory just to say that it is given by the grammar of English. For there may be competing grammars of English, and nothing we've said so far seems to determine that one rather than the other is the correct one. He thinks there is only one way in which the question can be answered: the right function is the one that speakers of English use to represent the sentences of their language. This of course commits us to the existence of something like an internal system of representation.

Objection 1

Why not take the objects of propositional attitudes to be propositions? Fodor thinks that, insofar as the internal representation system is a language, it should be possible in principle to offer a syntax and a semantics for it, and the hypothesis is in principle compatible with the claim that formulas in the language of thought express propositions or are related to propositions in pretty much the same sense in which sentences of natural language are related to propositions. In this sense, he is quite happy to say that propositional attitudes are mediated relations to propositions. However, his main worry is that, if propositions are understood as abstract entities, it's not clear how an organism can stand in a relation to a proposition except by standing in a causal or functional relation to some token formula which expresses the proposition.

Objection 2

It is conceivable that propositional attitudes are not relations to internal representations. Fodor thinks that this is true, but it's not damaging. He's offering an empirical account of the phenomenon, not an analysis. Contrast this with some physical theory: it is metaphysically possible that there is a world in which our best physical theory is false, but this doesn't mean that such theory is false in the actual world. To the more interesting question regarding the possibility of beings with different systems of representation, Fodor hopes that there will be translations between the systems of internal representation, so that expressions in formally different systems can express the same proposition.