

Frege's puzzle and some responses

Jeff Speaks
PHIL 93503

November 2, 2023

1 FREGE'S PUZZLES

One big question is what propositions are. But (as we've already seen in the discussion of 'sets of worlds' views) our answers to that question are constrained by our view about how fine-grained propositions are. One central way this second question gets raised is via language: under what conditions do a pair of sentences express the same proposition? Contemporary discussions of this are traceable to Frege's 1892 paper, 'On sense and reference.'

In a very famous passage, Frege says:

Equality gives rise to challenging questions which are not altogether easy to answer. Is it a relation? . . . $a=a$ and $a=b$ are obviously statements of differing cognitive value; $a=a$ holds a priori and, according to Kant, is to be labelled analytic, while statements of the form $a=b$ often contain very valuable extensions of our knowledge and cannot always be established a priori. The discovery that the rising sun is not new every morning, but always the same, was one of the most fertile astronomical discoveries. Even today the identification of a small planet or a comet is not always a matter of course. Now if we were to regard equality as a relation between that which the names ' a ' and ' b ' designate, it would seem that $a = b$ could not differ from $a = a$ (provided $a = b$ is true). A relation would thereby be expressed of a thing to itself, and indeed one in which each thing stands to itself but to no other thing.

One can think of this as a *reductio* argument. Suppose that the function of ' $a = b$ ' was just to predicate the relation of identity of a thing and itself. Then the function of ' $a = b$ ' would be the same as the function of ' $a = a$.' But these two claims are different in all sorts of ways: the first is a posteriori and can contain a valuable extension of our knowledge, whereas the second is trivial, analytic, and a priori. Frege's first puzzle: what explains this difference in informativeness?

Frege's answer: names must have a kind of meaning distinct from their reference. He called this kind of meaning their 'sense.'

There is a second, related puzzle which also gets called 'Frege's puzzle.' This does not come up explicitly in 'On sense and reference,' but is just below the surface. This is the puzzle of how substitution of names in attitude ascriptions can change their truth value. E.g.:

The ancients believed that Hesperus is Hesperus.

The ancients believed that Hesperus is Phosphorus.

Frege's second puzzle: what explains this difference in truth value?

Frege's answer: in attitude ascriptions names refer to their usual sense.

2 SOME FREGEAN THESES ABOUT PROPOSITIONS

What can Frege's puzzles tell us about propositions? Frege's solutions to the puzzles can be read as placing constraints on the conditions under which the propositions expressed by a pair of sentences are the same or different. And that constrains our theory of propositions.

Two important points about the puzzles. First: they have nothing special to do with identity. We could just as well have used

Hesperus is visible in the evening.

Phosphorus is visible in the evening.

Second point: arguably, they have nothing special to do with names. Kripke gives the example of 'furze' and 'gorse' which are terms for the same kind of bush. We could raise Frege's puzzles using

Every gorse is a gorse.

Every gorse is a furze.

So the arguments which seem to show that names have a kind of meaning which goes beyond the object designated also seem to show that predicates have a kind of meaning which goes beyond the property expressed.

In his later writings Frege suggested various versions of this criterion of difference for senses:

If it is possible to understand two sentences S and S^* while (after reflection) taking different attitudes toward their truth-values, then S and S^* differ in sense.

(Notice that this conditional can't be turned into a biconditional if there are some sentences which differ in meaning and no competent speaker could understand without seeing to be true/false.)

We can state this in proposition-talk like this:

Frege's criterion

If it is possible to understand two sentences S and S^* while (after reflection) taking different attitudes toward their truth-values in some context C , then S and S^* express different propositions in C .

This criterion corresponds to Frege's solution to the first version of Frege's puzzle.

This principle is closely related to a principle sometimes called the 'transparency of meaning.' This says that if someone understands a pair of expressions in a context, and in the context they have the same content, then on reflection the speaker will judge that they have the same content. Do these principles stand or fall together?

We can also state a condition for distinctness of propositions corresponding to the second version of Frege's puzzle:

Substitution

If two attitude ascriptions $\ulcorner A$ V 's that $S \urcorner$ and $\ulcorner A$ V 's that $S^* \urcorner$ can differ in truth-value in some context C , then S and S^* express different propositions in C .

We've already seen that there are cases involving substitution of coreferring names in attitude ascriptions which speakers take to involve a difference in the truth-value of the ascriptions. Proponents of broadly Fregean views of propositions typically take these intuitions at face value. Since some deny this, it will be useful to have a third rough thesis on the table:

Intuitions

Speaker intuitions about the truth-values of attitude ascriptions involving apparent substitution failures are correct.

Of course, one might only think that this holds in most cases, etc.

One can argue for Intuitions using two schemata that Kripke discusses:

Weak disquotation: If a competent speaker on reflection sincerely assents to ' S ', then that speaker believes that S .

Reverse disquotation: If a speaker believes that S then, on reflection, that speaker will sincerely assent to ' S .'

(Kripke never states the second principle; instead he combined this with weak disquotation to get the biconditional which he calls 'the strengthened disquotational principle.')

The apparent substitution failures in which we are interested are cases in which a speaker is disposed to accept one of the sentences, and not disposed to accept the other. So if our two disquotational principles are true, the first ascription will be true and the second false.

One can also use these principles, plus Substitution, to argue via conditional proof for Frege's criterion:

1. A understands each sentence and, upon reflection, thinks that S is true and S^* is false.
 2. $\ulcorner A$ believes that $S \urcorner$ is true. (1, weak disquotation)
 3. $\ulcorner A$ believes that $S^* \urcorner$ is false. (1, reverse disquotation)
-
- C. S and S^* differ in meaning. (2, 3, Substitution)

A classic Fregean view accepts all three of Frege's criterion, Substitution, and Intuitions. (Though there are Frege-adjacent views which reject some of these for some cases.) Anti-Fregeans typically reject Frege's criterion, but divide over whether they reject Substitution or Intuitions (or both).

3 SOME FREGEAN THEORIES OF PROPOSITIONS

Suppose we want to construct theories which respect all three of these Fregean principles. What might we take propositions to be?

3.1 Frege on sense

So far we've talked about how to tell when expressions do or do not have the same sense — but what sorts of things are senses? Here's Frege's explanation:

The reference of a proper name is the object itself which we designate by its means; the idea, which we have in that case, is wholly subjective; in between lies the sense, which is indeed no longer subjective like the idea, but is yet not the object itself. The following analogy will perhaps clarify these relationships. Somebody observes the Moon through a telescope. I compare the Moon itself to the reference; it is the object of the observation, mediated by the real image projected by the object glass in the interior of the telescope, and by the retinal image of the observer. The former I compare to the sense, the latter is like the idea or experience. The optical image in the telescope is indeed one-sided and dependent upon the standpoint of observation; but it is still objective, inasmuch as it can be used by several observers. At any rate it could be arranged for several to use it simultaneously. But each one would have his own retinal image. ('On sense and reference,' 30)

Senses are abstracta which somehow determine a reference. Notoriously, Frege said few non-metaphorical things about what senses are and how they determine a reference.

3.2 Descriptivism

Often when Frege explains the sense of a proper name, he does so using a definite description. This view might help with Frege's puzzles if the meanings of coreferential names were given by distinct co-designating definite descriptions. So, e.g., the meaning of 'Hesperus' might be 'the brightest object in the evening sky' and analogously for 'Phosphorus.'

This view was criticized by Kripke in *Naming and Necessity*. Some of the criticisms:

- Names and the corresponding descriptions aren't substitutable in modal contexts. (the modal argument)
- Ditto for some epistemic contexts, like 'it is a priori that.' (the epistemic argument)
- Often the descriptions associated with names by speakers don't even pick out the same thing as the name. Examples of 'Einstein,' 'Gödel.' (the semantic argument)

3.3 Two-dimensionalism

Last time we discussed the idea that we can associate names with primary intensions — functions from worlds considered as actual to extensions — as well as secondary intensions. (If names are rigid, the latter will be constant functions.)

This can be thought of as a kind of successor theory to descriptivism. Some two-dimensionalists want primary intensions to be stutable as descriptions, others don't. But on most views primary intensions are like descriptions in being 'ways of determining a reference' which for a pair of names might coincide at the actual world but diverge at other worlds considered as actual.

Why the primary intensions of ‘Hesperus’ and ‘Phosphorus’ plausibly diverge, giving us a kind of difference in meaning.

How should the two-dimensionalist respond to Kripke’s arguments?

3.4 *Impossible worlds semantics*

As discussed, this has the potential to make plenty of distinctions in meaning. Could also be combined with two-dimensionalism in various ways.

4 KRIPKÉ’S PUZZLE

4.1 *The original puzzle*

Kripke’s puzzle is an attempt to undercut some of the motivations for our Fregean package of views. It employs the two disquotational principles already mentioned, along with two more:

Translation: if a sentence expresses a truth in one language, then its translation into another language also expresses a truth (in that language).

Contradiction: if you have a pair of contradictory beliefs, it is always possible to discover this a priori.

Puzzling Pierre: accepts

Londres est jolie.

So, by Weak Disquotation and Translation, it follows that

Pierre believes that London is pretty.

But Pierre also accepts

London is not pretty.

But Weak Disquotation then we have

Pierre believes that London is not pretty.

So by Contradiction he is in position to discover that he has a contradictory belief a priori. But he isn’t. So one of Weak Disquotation, Translation, and Contradiction is false.

Second version: start as before, but now note that Pierre rejects

London is pretty.

So, by Reverse Disquotation, we have

It is not the case that Pierre believes that London is pretty.

which is a contradiction. So at least one of Weak Disquotation, Reverse Disquotation, and Translation is false.

Translation looks plausible. Further, we can construct a case without Translation: Peter and Paderewski.

Kripke does not himself take a stand on how to resolve these paradoxes. A common anti-Fregean take: resolve the first version of the puzzle by rejecting Contradiction, and the second by rejecting Reverse Disquotation.

How does this undercut the attempt to argue from Frege's puzzle to Fregean views of propositions?

One might argue as follows: 'Lois believes that Clark cannot fly and believes that Superman can. But Lois is not irrational, couldn't discover the conflict a priori, etc. So, the two propositions can't really be contradictories of each other.' But this employs Contradiction.

Or one might argue as follows: 'Lois accepts the sentence 'Superman flies', and so believes that Superman flies. But of course (after as much reflection as you like) she explicitly rejects 'Clark flies', so she does not believe that Clark flies. So, the proposition that Clark flies must be distinct from the proposition that Superman flies. But this uses Reverse Disquotation.

4.2 *Other parallel arguments*

Salmon gives a parallel argument: Sasha and 'catsup' and 'ketchup.' It seems that you can run a similar argument for any alleged synonyms not recognized as such. One can see this argument (as well as Kripke's) as targeting the transparency of meaning. Can also use them to argue against Frege's criterion and against the conjunction of Substitution and Intuitions.

4.3 *Fregean replies*

Fregeans tend not to be impressed by this line of argument. Some Fregean replies:

1. Pierre, Sasha, et. al. don't really understand the relevant expressions. But: they seem to understand them as well as we understand lots of expressions of our language.
2. 'Londres'/'London' don't mean the same thing, so the sentences aren't really meaning-preserving translations of each other. Ditto for 'catsup'/'ketchup' and the two tokens of 'Paderewski.' But: surely synonymy is possible!
3. Maybe the expressions do mean the same thing (even out of the mouths of Pierre and Sasha) but the relationship between public language and mental content is more indirect – Pierre's 'London' beliefs do differ in content from his 'Londres' beliefs, but it's hard to express this difference in language. This seems to involve giving up on Frege's criterion and one of Substitution or Intuitions — gets a little murky at this point how to draw the line between Fregean and non-Fregean views.

4. The relevant pairs of expressions might mean the same thing for some people, but they don't mean the same for the relevant language users. The meanings of names and other expressions typically vary between speakers.

Lots of Fregeans go for #4. And it arguably falls out of at least some versions of two-dimensionalism that, while secondary intensions may be stable across (most) speakers, primary intensions will be highly variable. This might look useful for explaining various linguistic phenomena; we explain some in terms of the stable semantic features, and others in terms of the more individualistic ones.

4.4 *Explaining substitution success*

The plausible idea is surely that we explain attitude ascriptions in terms of some combination of the two. The question is how to do this.

It is instructive to begin with a theory that (to my knowledge) no one has ever proposed, which requires for the truth of an attitude ascription that the complement in the context have the same primary and secondary intension as the belief of the subject of the ascription.

The reason why no one has ever proposed this is that it is way too strong. Some examples:

Daniel believes that I make good cocktails.

Aristotle believed that Plato was pretty smart.

So how should the account be relaxed? Maybe: 'similar enough' primary intensions. But what counts as 'similar enough' will depend on context. We can give a context in which the following seem to differ in truth value:

Cynthia knew that Ted Warfield is Ted Warfield.

Cynthia knew that Ted Warfield is Fritz Warfield.

If they differ in truth-value, then Cynthia's 'Fritz' primary intension must be similar enough (in this context) to my 'Fritz' intension but not similar enough to my 'Ted' intension. This has to be possible in some contexts even if my 'Ted' and 'Fritz' primary intensions are extremely similar to each other, and even if both are much more similar to Cynthia's 'Fritz' intension than anything in the Aristotle/Plato case.

So what we need is a theory which tells us, for arbitrary contexts and ascriptions, what the relevant relation between primary intensions is.

In general, there are two aspects of giving a Fregean solution to Frege's (second) puzzle: (i) providing fine-grained enough entities, and (ii) explaining how those entities explain the truth conditions of the relevant ascriptions. (i) can definitely be done — arguably the jury is out on (ii).

Could the Fregean reply by letting the relevant context-sensitive relation between primary intensions be primitive? They could say something like:

'A believes that S' is true in C iff (i) the secondary intension of S in C is the secondary intension of one of A's beliefs and (ii) the primary intension of S in C stands in the similar-enough-in-the-right-ways-for-C relation to A's belief.

where the underlined relation is not further explained.

People have different views on whether this kind of thing is ok. In favor: we must have some implicit grasp of such a relation; otherwise where does our knowledge of the truth conditions of attitude ascriptions come from? Against: worries about whether this sort of thing can explain our semantic competence. Compare:

‘A believes that S’ is true in C iff S stands in the belief-matching-in-C relation to one of A’s beliefs.

5 REJECTING INTUITIONS

One kind of anti-Fregean thinks that speaker intuitions about the truth values of ascriptions in Frege cases are incorrect.

An obvious question for the anti-Fregean who rejects Intuitions: why do speakers systematically make mistakes about the truth-values of these ascriptions?

5.1 *The appeal to pragmatics*

One answer: they are confusing the semantic content of the ascription with the information pragmatically conveyed by its utterance.

Everyone agrees that this distinction exists, and almost everyone agrees on some paradigm examples of it.

How would it apply to the present case?

Salmon (1986) argues that belief is a three place relation between a subject, a proposition, and a way of encountering or thinking of that proposition. Belief ascriptions express existential generalizations of this three place relation. But they pragmatically convey information about how the subject in question encounters the proposition. So, e.g.

Lois believes that Clark Kent is Superman.

semantically expresses the same proposition as

Lois believes that Superman is Superman.

but the former conveys (perhaps) that Lois assents to this proposition while encountering it via the sentence ‘Clark Kent is Superman.’ This proposition is false; that’s why we take the ascription (incorrectly) to be false.

Soames (2002) argues that utterances of ascriptions convey (and assert) descriptively enriched propositions, exploiting associations of expressions with descriptive information in the conversation. Other variants of this general sort of view are available.

Core of the strategy: save the idea that speaker intuitions are correct about *something*, even if not about the propositions expressed by the ascriptions.

Some objections:

1. Mysterious pragmatic mechanism; can occur in thought as well as in conversation.
2. Do we really get that confused about semantics vs. pragmatics? It often seems to be clear enough.
3. Terminological variant on the Fregean view – just a fight about how to use ‘semantics’ and ‘pragmatics.’

5.2 *The appeal to heuristics*

Williamson, unlike Salmon and Soames, tries to give an error theory for Frege cases which explains speaker intuitions without trying to make them right about something. The idea that we use heuristics — which in some cases may approximate the disquotational principles — in assessing the truth of attitude ascriptions. In Frege cases we can expect those to lead us astray.

Williamson’s challenge to the Russellian: if heuristics can explain our Fregean intuitions, why not just go for intensionalism?

Reply to Williamson: we don’t just need a theory of why certain attitude ascriptions involving mathematical belief seem to have certain truth conditions; we also need an account of what is going on in (apparent) mathematical cognition. Seems little alternative to the Stalnaker story. The Russellian seems better off here.

6 REJECTING SUBSTITUTION

A different strategy for the anti-Fregean is to reject Substitution. One way to do this: adopt the view that attitude ascriptions are context-sensitive in a certain way.

An argument that attitude ascriptions are context-sensitive: Schiffer’s example of Shorty and ‘She knows that I limp.’

One way to do this: take attitude ascriptions to express something like the propositions Salmon takes them to pragmatically convey. On this view ‘believes’ expresses a three place relation, whose third relatum is an ‘unarticulated constituent.’ The example of ‘ready.’

Another way to do this: ‘believes’ is an indexical, so that it expresses different two-place relations in different contexts. Why this view has to posit a lot of context shifting.

REFERENCES

- Nathan Salmon, 1986. *Frege’s Puzzle*. Atascadero, CA: Ridgeview Publishing Company.
- Scott Soames, 2002. *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*. Oxford: Oxford University Press.