

# 13. Frege on sense and reference

Martín Abreu Zavaleta

July 28, 2015

We have been talking about identity for around a week, focusing on what makes people who exist at different times be one and the same person. Our treatment of such issues so far has been of a *metaphysical* nature: we have tried to find out in virtue of what a person is identical to some person at a later point in time. Another feature of our discussion is that it has focused on identity *through time*.

However, identity through time is not the only way in which identity can be puzzling. In everyday life, we also talk about identity at a given time: we say things like “ $2+2=4$ ” or, discover that what we took to be two different people is in fact a single person—we’ll see some examples in a moment.

Like Locke and the other philosophers whose work we discussed, Frege was also concerned with identity. However, he was not so interested in the metaphysical questions concerning when objects are identical to each other (or at least not in the paper that you read). He was more interested in identity from a *linguistic* and *logical* point of view. This should be hardly surprising: Frege was the inventor of modern logic (what we now call classical logic), and as such, his research focused on the properties of certain formal languages and logical concepts.

His questions concerning identity is not so much ‘what makes two things identical to each other?’ but ‘*What do we express, or what do we mean when we use identity statements?*’ We will now address identity from this linguistic perspective, which will ultimately lead us to study the nature of names and other referring expressions.

Frege’s work has been fruitful in many ways. Besides inventing modern logic, he could reasonably be credited with the discovery and original formulation of many problems that are still studied by contemporary philosophers of language. His paper ‘On Sense and Reference’, originally published in 1892, has influenced generations of philosophers and, for better or worse, framed some foundational questions in the philosophy of language.

## 1 Cognitive significance

Frege’s whole paper is concerned with identity statements. In particular, he is concerned with what we now call ‘Frege’s puzzle of cognitive significance’ or ‘Frege’s puzzle’ for short, which I’ll introduce by means of an example.

Suppose over the course of several conversations, I tell you that I have a friend called Carla, and I tell you several things about her. I tell you that she does philosophy of mind, that she used to live in New Jersey, that she is very smart, etc. In some other conversations, I talk to you about my friend Smith. I tell you that Maria lives in North Carolina, that she used to work in Arizona, and so on.

Now suppose one day in the middle of a conversation you ask me something concerning Carla and Maria. For instance, you ask me if Carla and Maria will come to my party, after which we have the following conversation:

- M: Carla said she was definitely coming.
- You: What about Maria?
- M: Well, if Carla is coming then Maria is coming too: Carla is the same person as Maria.

When I say ‘Carla is the same person as Maria’, or simply ‘Carla is Maria’, you learn something that you didn’t know before, namely, that Carla is Maria. However, if I had instead told you ‘Carla is Carla’, you wouldn’t have learnt anything new: you already knew that Carla was the same person as Carla.

More generally, if ‘a’ and ‘b’ are names, then a sentence of the form ‘a=b’ can be informative, but a sentence of the form ‘a=a’ usually isn’t. This is what Frege means when he says “a=a and a=b are obviously statements of differing cognitive value” (p. 209). A statement of the form ‘a=b’ is potentially informative, like when I say that Maria is Carla, but a sentence of the form ‘a=a’ is never informative (or, at best, could only be informative to someone who didn’t know that some object is identical to itself). This difference in informativity, or as Frege calls it, *cognitive value*, is our starting data point:

**Difference in cognitive significance:** Sentences of the form ‘a=b’ and sentences of the form ‘a=a’ differ from each other in cognitive value.

Now let’s pay some more attention to what this means, exactly.

Above I briefly mentioned that ‘Carla=Carla’ could be informative for at least some people, namely someone who either didn’t know that everything is identical to itself, or someone who didn’t know that Carla in particular is identical to herself. If this is true, then it seems that statements of the form ‘a=a’ *can* be informative after all. Perhaps a person for whom ‘Carla=Carla’ was informative would not be specially smart or rational, but it would still be informative for him. So the difference in cognitive value can’t be just that one kind of sentence *can* be informative for someone while the other one *can not*. Yet there seems to be a difference between ‘a=a’ and ‘a=b’. How should we understand that difference?

Frege uses a technical term to clarify what he means. He says “a=a holds a priori and, according to Kant, is to be labeled analytic, while statements of the form a=b often contain very valuable extensions of our knowledge and cannot always be established a priori.” (p. 209) If we are to understand what Frege means, we better know what ‘a priori’ and ‘analytic’ mean.

Aside: a priori/a posteriori, analytic/synthetic

‘A priori’, ‘a posteriori’, ‘analytic’, and ‘synthetic’, are fairly old technical terms. I’m not sure if they were introduced by Kant, but Kant definitely used them. Kant meant something different from what we now mean by them, so let’s just describe their contemporary use. As a preliminary, these four expressions are used to describe judgments or claims.

The distinction between a priori and a posteriori claims has to do with the way in which we can know that the claim is true. Roughly, a claim is a priori if we can come to know it without requiring empirical evidence (evidence e.g. from our senses). For instance, the claim that everything is

identical to itself is a priori, because we don't need empirical evidence in order to come to know it. A claim is a posteriori if we can only know it by means of some empirical evidence. For instance, the claim that there are planets, or that sheep are mammals, can only be known by means of empirical evidence.

Notice that a priori claims *can* be known through empirical evidence. For instance, you may come to know it because your math teacher told you, or because you read it in a book. Since your math teacher and the book are reliable sources of evidence, you come to know that everything is identical to itself when you hear it from them. The point, however, is that, at least in principle, you didn't have to learn it that way *necessarily*. If we had put you in an isolated room without light or any sensorial stimulation, you could have come to learn it by reasoning alone. Perhaps even the development of your reasoning abilities required you to engage in some sort of empirical investigation, but that's besides the point. Even if there is some *causal* connection between your knowledge of a claim and some piece of empirical knowledge, a claim is a priori as long as you can come to know it without any *evidential* connection between the claim and the piece of empirical knowledge.

The distinction between analytic and synthetic claims has to do with what makes them true. Analytic claims are supposed to be true merely in virtue of their meaning. For instance, because 'vixen' pretty much just means 'female fox', the sentence 'every vixen is a female fox' is supposed to be analytic. It is true because of what 'vixen' and 'female fox' mean (together with the meanings of the other words in the sentence).

Synthetic claims are not supposed to be true in virtue of their meaning. Instead, they are supposed to be true in virtue of how the world is. For instance, if 'some swans are black' is true, it is not because of what 'swan' means, but because of the color of certain swans.

Note that even synthetic claims are true partly in virtue of their meanings: if 'some swans are black' is true, this is partly because of what that sentence means. Namely, it means that some swans are black. If it instead meant that  $2+2=5$ , then it wouldn't be true. The crucial difference is that, unlike analytic claims, synthetic claims are not true *merely* because of what they mean. The distinction between analytic and synthetic judgments was criticized during the 60s and today very few philosophers use it to make substantive philosophical points.

## Back to cognitive significance

So perhaps the best way to understand the difference in cognitive value between 'a=a' and 'a=b' is by saying that the first is a priori but the second one is not. Notice, however, that we could characterize the difference that Frege is after even without using the notion of apriority: we could just point out that it is possible that 'a=b' but not 'a=a' is informative *for someone*. In that case, the data point will be that *some* people can learn something new when they are told that a=b but not that a=a. For the rest of our discussion, we will continue to understand the difference in cognitive value between a=a and a=b in terms of apriority, since this seems closer to what Frege meant.

## 2 The puzzle of cognitive significance

Frege's claim about the difference in cognitive significance between 'a=a' and 'a=b' seems intuitively compelling. However, it leads to a puzzle: what do these two sentences express? On one hand, we could take each sentence to establish a relation between the thing named by the first name

and the thing named by the second one. But if this is so, then 'a=a' and 'a=b' would express exactly the same thing, provided that they are true. But if they express the same thing, how could one be informative but not the other?

Frege briefly considers an alternative that he endorsed in previous work, but which he rejects in this paper. The alternative is that 'a=a' and 'a=b' don't express a relation between objects, but between the symbols or signs that are part of the sentence. That is, they say something like 'the name "a" denotes the same thing as the name "a"' and 'the name "a" expresses the same thing as the name "b"', respectively. Call this the *metalinguistic view*.

Frege rejects this last view because he thinks that it fails to account for an important feature of sentences like 'a=b'. This kind of sentence gives us *genuine knowledge* about its subject matter. When I told you that Carla was Maria, you learnt something *about Carla*. Of course, you also learnt something about my use of the names 'Carla' and 'Maria', but that's not all you learn.

Frege's argument for rejecting the metalinguistic view is not all that clear. One way of reconstructing it is by taking seriously his remark that the fact that two signs refer to the same thing is arbitrary. What does he mean? Well, think about what happens when someone names a baby or a pet. She chooses a name for the baby or the pet, but she does so only on the basis of the names that she likes, whether they sound nice, maybe whether they are the names of people she respects, etc. All these features are arbitrary from the perspective of the properties of the newly named object. There is nothing in me, for instance, that would require me to be connected to the name 'Martín'.

Similarity with 'Carla' and 'Maria'. That these two names happen to name the same person is just a coincidence, or a matter of convention. It has nothing to do with the properties that Carla has, or with how Carla is. But *genuine or proper* knowledge, the kind knowledge that statements like 'Carla is Maria' are supposed to give us, must be knowledge about Carla's properties, or how Carla is. Thus, identity statements of the form 'a=b' can't be statements that merely relate two names by saying that they have the same denotation.

If this doesn't sound convincing, think about what happens when we introduce a friend to someone else. For instance, I may introduce Carla to you by saying 'she is Carla'. If the metalinguistic view is right, all I would have said is that the expression 'she' as I used it in my utterance denoted the same thing as the name 'Carla'. But at the very least, we would have expected my introduction to give you a piece of information *about* Carla, even if that piece of information is just that her name is 'Carla'. No such information is provided if the metalinguistic view is correct.

Frege's argument points us in the direction of a different kind of argument. Suppose that 'a=b' truly expresses a relation between the names 'a' and 'b'. Then, in the absence of a special explanation why 'a' and 'b' don't denote their referents in identity sentences, we should assume that names never *contribute* their referents to any sentence in which they occur. But then what should we make of simple predications like 'Carla is a woman'? Ordinarily, we would think that a sentence like that says that Carla has a certain property, namely, being a woman. In the absence of a special explanation about the behavior of names in identity sentences, the metalinguistic view assigns an odd meaning to 'Carla is a woman'. Perhaps something like *the name 'Carla' overlaps with the property of being a woman* or *the name 'Carla' is the name of a woman*. But if anything, these claims are *entailed* by what I said. What I said was the simple thought that Carla is a woman, or, if you want to make it sound more complicated, that Carla has the property of being a woman.

So it seems the metalinguistic view (or, at least, this version of it) is not very promising. At this point, Frege introduces his new solution to the puzzle.

### 3 Modes of presentation

Frege's solution to his own puzzle requires him to postulate two dimensions in the meaning of a proper name. On one hand, names have a *referent*, the thing that they denote, if any. On the other, names also have a *sense* or *mode of presentation*. The idea is that identity sentences are relations between objects, but the names carry more information with them. Because of the modes of presentation of 'a' and 'b' are different, the identity sentences 'a=a' and 'a=b' carry different information: the first is not informative because 'a' presents its denotation in the same way as 'a'; the second is informative because 'a' and 'b' present their denotation in different ways. This is what Frege means when he says "A difference [in cognitive value] can only arise if the difference between the signs corresponds to a difference in the mode of presentation of that which is designated" (p. 210)

Frege elaborates on the notion of a mode of presentation by using an example:

Let  $a, b, c$  be the lines connecting the vertices of a triangle with the midpoints of the opposite sides. The point of intersection of  $a$  and  $b$  is then the same as the point of intersection of  $b$  and  $c$ . So we have different designations for the same point, and these names ('Point of intersection of  $a$  and  $b$ ', 'Point of intersection of  $b$  and  $c$ ') likewise indicate the mode of presentation; and hence the statement contains true knowledge.

It is natural, now, to think of there being connected with a sign (name, combination of words, letter), besides that to which the sign refers, which may be called the referent of the sign, also what I would like to call the *sense* of the sign, wherein the mode of presentation is contained. In our example, accordingly, the referents of the expressions 'the point of intersection of  $a$  and  $b$ ' and 'the point of intersection of  $b$  and  $c$ ' would be the same, but not their senses. (p. 210)