Marco Ruffino (*)

Frege's Puzzle: Can we pose it on Frege's Behalf? [1]

Resumen: *En este artículo reviso primeramente los elementos principales del así denominado puzzle de Frege, y argumento que hay algo extraño en el argumento que Frege edifica sobre este: primeramente, Frege rechaza una posible hipótesis respecto de la naturaleza de la identidad para hacer plausible la distinción entre sentido y referencia. Sin embargo, después de que la distinción es formulada, la hipótesis rechazada es, después de todo, la única compatible con ella. A continuación discuto la posición de Glezakos (2009) respecto del puzzle. Argumento que, aun cuando ella apunta algo muy importante, no hemos de aceptar su conclusión de que no hay un puzzle que puede formularse en términos neutrales.*

Palabras claves: *Puzzle de Frege. Identidad. Valor cognitivo. Nombres. Demostrativos.*

Abstract: *In this paper I first review the main elements of the so-called Frege's Puzzle, and argue that there is something odd in the argument that Frege builds based on it: Frege first rejects a possible hypothesis for the nature of identity in order to make plausible the distinction between sense and reference. But, after the distinction is made, the rejected hypothesis is the only one compatible with it after all. Next, I discuss Glezakos' (2009) position regarding the Puzzle. I argue that, although she does point out something quite important, we do not have to accept her conclusion that there is no puzzle that can be formulated in neutral terms.*

Key words: *Frege's Puzzle. Identity. Cognitive value. Names. Demonstratives.*

The so-called *Frege's Puzzle* in the literature was not only the motivating piece behind Frege's distinction between sense and reference, but has also triggered an intense discussion concerning the notion of cognitive content of language (in particular, of identity statements). Glezakos (2009) raises an important methodological issue about Frege's original formulation of the Puzzle: according to her, if carefully considered, it is either not a puzzle after all, or, if it is a puzzle, it can only be formulated under the assumption of the distinction between sense and reference (which is meant to be a solution for it, and not a starting point). In this paper I shall first review Frege's formulation of the Puzzle. I will try to show that there is another problem for its formulation besides the one pointed out by Glezakos. Then, I will argue that, although Glezakos raises an interesting challenge to Frege's original formulation, we can re-formulate the Puzzle on Frege's behalf, retaining the essential aspects of Frege's intention.

1. Frege's Puzzle and Senses: The Standard Story

In the opening sections of *SuB* Frege seems to be dealing with two different groups of questions, and it is important to be clear about the difference between them. The first (opening) group of questions are: (i) is identity a relation? And (ii) if so, what are its *relata*? The second group of questions is: (iii) what is the explanation for the cognitive difference between the statements $a=a$ and $a=b$ (in case they are both true)? and (iv) why is the first analytic and known a priori, while the second, if true, "cannot always be established a

priori" (*SuB*, p. 25)? Questions of the first group are primarily metaphysical, and concern the very nature of identity *qua* thing. Questions of the second group are epistemic and concerned with the cognitive content of identity-sentences. Part of Frege's initial arguments in *SuB* is that some possible answers to the metaphysical questions do not yield satisfactory answers to the epistemic questions. One should notice that, although in most examples that Frege uses along the text *a=b* are identities that can only be known only a posteriori (like, e.g., *Hesperus is Phosphorus*), this is not a necessary feature of the phenomenon of cognitive difference that he is pointing at. For *a=b* might be a non-trivial identity of arithmetic, in which *a* and *b* refer to the same number but as a result of different operations, and hence might be knowable a priori as much as *a=a*, but nevertheless express something different from the latter. This leads us to suppose that the correct interpretation of Frege's remark is not that we should look for an explanation of the fact that *a=a* is a priori while *a=b* is a posteriori (in case they are both true) but that any explanation of the nature of identity (i. e., any answer to questions of the first group) should leave open the possibility of there being cases of *a=b* that are knowable only a posteriori, although *a=a* is, presumably, always knowable a priori.

I think that what is normally described in the literature as the Puzzle has to do more with the second group of questions. But both groups of questions are proposed as a motivation for Frege's distinction between sense and reference, and it is almost a universal consensus that the latter is is meant as a solution for them.

In the opening paragraphs of *SuB* Frege first considers two possible answers to (ii), and discharges both because they fail to yield plausible answers to (iv). The first is that the *relata* are "that which the names '*a*' and '*b*' designate [*bedeuten*]" (*SuB*, p. 25), i. e., the objects that '*a*' and '*b*' stand for. (Notice that this and any other answer to (ii) presuppose an affirmative answer to (i), for obviously if one does not consider identity as a relation, there is no point in asking for the *relata*.) Let us call this the *Object-Alternative* (**OA**). But if **OA** is correct, then, according to Frege, *a=a* and *a=b* say exactly the same, i.e., that the object is related to itself, and there is no way of leaving open the possibility that *a=b* might express something different from *a=a* (and, therefore, of being knowable only a posteriori). Hence, Frege seems to imply, there is no answer to (iv). As we know, Frege considers as second possible candidate for *relata* the expressions *a* and *b* themselves. Let us call the the *Name-Alternative* (**NA**). If **NA** is correct, supposedly identity would amount to something like *co-designation* or *co-reference*. This is possible and coherent, but Frege discharges **NA** because, according to him, since co-referentiality is a matter of convention, the solution would treat any true identity like *a=b* as "no longer be concerned with the subject matter, but only with its mode of designation" (26). In this sense, all true identities would express linguistic facts, but no astronomical, mathematical, etc., facts. Frege seems to suggest that this goes against the natural view that true identities do express non-linguistic facts. And indeed this is so. Suppose that in a medical conference some scientist states a true identity that we all have been long waiting for (and that will have a deep impact on our lives) such as 'the drug that completely cures any form of cancer is such-and-such'. Despite of our joy and relief, **NA** would lead us to conclude that the scientist did nothing but establishing a new linguistic rule (which is, moreover, arbitrary). This seems deeply wrong (and unfair to the scientist's intention and achievement).

As we know, at the end of his argument, Frege says that a difference in cognitive value between *a=a* and *a=b* can only exists if there is, corresponding to the difference in the signs *a* and *b*, different modes of presentation ("*Art des Gegebenseins*") of the common reference, and the mode of presentation is what he will call the sense ('*Sinn*'). Hence, apparently the answer to (iii) is that *a=a* and *a=b* have different senses, since *a* and *b* have different senses.

2. A Puzzle About Frege's Puzzle

Despite appearances, it is not clear how exactly Frege uses the distinction between sense

and reference to solve questions (ii)-(iv). (We may assume, I think, that his answer to (i) is affirmative.) Curiously enough, after introducing the distinction Frege does not go back to question (ii), and does not tell us what the *relata* of identity are after all. But we can perhaps infer that. They cannot be the senses, for one of the conclusions of his discussion is that, since we have in *a=b* two different names flanking the identity sign, we must have two different modes of presentation and, hence, identity cannot hold of them. I agree with Perry (2009) that the *relata* must be the reference of *a* and *b*. Actually, after the distinction between sense and reference, it would seem that the question regarding the nature of identity has to be unfolded into two separate questions: the nature of the reference of the identity-sign, and the nature of the sense of the identity sign. Only the reference of the identity-sign is a relation, i .e., something that associates truth or falsity to pairs of objects. The sense of the identity-sign does not associate truth or falsity to its arguments (two singular senses), but only senses (i. e., complete thoughts). Hence, it is not a relation, but a binary function from pairs of singular senses to thoughts. Something similar could be said of the identity sign itself: it is not a relation, but a binary function from singular terms to sentences. Hence, if the answer to (i) is affirmative (i. e., identity is a relation), the only alternative is to consider it as a relation that has pairs of objects as arguments, and is true only of the pair of an object taken with itself. I.e., we are back to **OA**.

But if this is correct, there seems to be something strange in Frege's dialectic. For, remember, one of the steps that lead him to establish the distinction between sense and reference is the denial of **OA**. He first excludes it in order to make plausible a distinction that, once established, leads to the conclusion that **OA** is the only alternative available after all. How can that be?

We can only conclude the following: both the denial and the acceptance of **OA** are compatible with the distinction between sense and reference. Hence, the distinction is not needed as an answer to (ii) (or to (i)), but only to (iii) and (iv).

3. Is Frege's Puzzle Puzzling?

Glezakos argues that, if carefully considered, Frege's Puzzle either cannot be formulated (i. e., is not a puzzle at all) or, if it can be formulated, its formulation requires the very conceptual apparatus introduced by the distinction between sense and reference, which was supposed to be a solution for the Puzzle.

The dilemma comes from taking a closer look into the general structure of Frege's Puzzle and the presuppositions behind it. Frege assumes sentences of the form *a=a* and *a=b* always differ in cognitive content. But this is highly schematic. How can we know whether an identity containing names has the form *a=a* or *a=b*? As Glezakos says, "The most natural account of what determines an identity sentence's form involves the notion of sameness of name. Without making any theoretical commitments, we can say that a sentence has the form *a=a* when the same name flanks the identity sign, and *a=b* when distinct names flank the identity sign" (p. 203).

According to her, this is neutral enough so as not to commit us to any sort of theory from the outset. But this raises the question of *name-individuation* (and here she is inspired by Kaplan (1990) in asking for the basis of word individuation.). So, how can we know that, e.g., *Socrates is Socrates* has the form *a=a* or *a=b*?[2]

Now Glezakos points out two distinct ways of distinguishing names that one can find in Frege's writings. The first one appears in a famous footnote of *SuB* in which he seems to suggest that an ordinary name like *Aristotle* might have different senses attached to it since "opinions as to the sense may differ", but that would be tolerable "so long as the *Bedeutung* remains the same" (*SuB* 27, note B). This suggests, according to her, one characterization of names as "a particular sign for a particular referent" (2009, 204). Presumably this criterion of name-identity would consider two names as the same iff they are tokens of the same type and have the same referent. She concludes that, under this criterion, "there will be no epistemic divide between sentences of the form *a=a* and *a=b*" (ibid.). She takes as basis for this claim the Paderewski-like

cases famously discussed by Kripke (1979) (i. e., one might hear about Paderewski as the musician, and later as a politician, and could perhaps conclude that there are two names *Paderewski*, and hence be surprised to discover that *Paderewski is Paderewski* is true). But since the epistemic difference is the basis for the puzzment of Frege's Puzzle, it follows, according to her, that there is no puzzle.[3]

Another criterion for name-distinction in Frege's writings detected by Glezakos is a bit more radical, and employs the notion of sense. In a famous passage of "*Der Gedanke*" (1918-19) about the name *Dr. Lauben*, Frege explicitly says regarding two persons that attach different senses to it that "[they] do not speak the same language, although they do in fact designate the same man with this name; for they do not know that they are doing so" (65). That is to say, since *Dr. Lauben* is taken with two different senses by the two speakers, the same expression corresponds to two different names. This would yield a way of deciding whether an identity is of the form *a=a* or *a=b*. But, as she points out, "[i]f a name is in part individuated by the *Sinn* associated with it, then, in order to be puzzled, one would need to be committed to *Sinne* and their role in determining name identity" (2009, 206).

I have two objections to Glezakos' points. First, if I understand it correctly, her claim is that the phenomenon that Frege wants to call our attention for (differences in cognitive content) cannot be produced in any form of identity unless the distinction between sense and reference is already established. She does not consider, however, whether there could be any difference if, instead of ordinary proper names, we have other forms of singular terms flanking the identity sign. (This might be suggested by Frege's procedure, since he himself does not raise the question.) Identity between ordinary proper names is one particular form that Frege uses to motivate his distinction between sense and reference. But it might not be the only form of identity available to generate the phenomenon.

As I see it, if we can formulate one single pair of identities that exhibit the aspect of cognitive difference combined with co-referentiality of the proper names involved, this would already be enough to generate a puzzle that calls for something like Frege's distinction. We could escape this dilemma if we had a situation in which, without assuming beforehand the notion of sense, we can tell whether we have the same or different names flanking the identity sign. Sameness of reference is not enough for that, as Glezakos argues. And differences in sense is not open. But maybe we can have a way of constructing identities in which the names refer to the same object, and the criterion of identification is not given simply by the words themselves, but by some other extra-linguistic element. I think we can find that in identities involving demonstratives. Consider the following true identities between demonstratives:

That[A]=That[A] and That[A]=That[B]

where *A* and *B* are accompanying demonstrations (i. e., a gesture like pointing or any sort of extra-linguistic elements that have the same function) of the same object. In his famous investigation about the properties of demonstratives, Kaplan says that "demonstratives are incomplete expressions which must be completed by a demonstration (type)" (1989, 527). That means that the sameness of names (broadly conceived) in the case of demonstratives is given neither by the same form of the linguistic expression (since it is the same in *that[A]* as in *that[B]*) nor by the same object referred, but by something external (the demonstration). So, we do know immediately that in *That[A]=That[A]* we have the same name flanking both sides of identity (because the demonstration is the same), while in *That[A]=That[B]* we have two distinct names (because the demonstrations are different). Demonstrations, *qua* extra linguistic devices, do the job of word-individuation in the case of demonstratives, and provide an adequate context to produce the kind of phenomenon of cognitive value that Frege was after. He would not formulate it this way since he tends to think of demonstratives and indexicals in general more like a sort of pathology of natural language. But we can re-phrase the puzzle on Frege's behalf in a theory-neutral way.[4]

My second objection is the following: it seems to me that it ought to be possible to make

a distinction (not made by Glezakos) between recognizing that two sentences of the forms *a=a* and *a=b* express different cognitive contents on the one hand, and recognizing that particular sentences have the form *a=a* or *a=b* on the other hand. The point Frege is making is (or so we can re-phrase it) that, given that we have already decided that two sentences are of the form *a=a* and *a=b*, then we see a difference in cognitive content between them, and we see that the former is knowable a priori, while the latter might not be so (or, if knowable a priori, it is knowable by a different process than the former). This is compatible with saying that the process by means of which we discover whether the sentence has the form *a=a* or *a=b* might be much more complex than simply knowing that both forms have different cognitive values. Consider *I+I=II*. Here is a paradigmatic example of something that can be known a priori. However, there might be some doubt as to whether the second occurrence of *I* is the same roman numeral as the first, or whether it is the indexical *I* (used for myself), and some sort of empirical investigation might be necessary to be clear about the form of the sentence. This does not destroy the credentials of the sentence as knowable a priori.

I do think that Glezakos' challenge is important because it helps to better understand Frege's Puzzle. In particular, it helps to understand what is essential for its formulation and what is simply a feature of the particular way Frege chooses to present it. But it does not show that, as she is tempted to conclude, something to the same effect as Frege's Puzzle cannot be posed in a theory-neutral way.

Notes

1. I first became interested in Frege's Puzzle after a long discussion with David Kaplan during my UCLA years, to whom I express my gratitude. Thanks also to Manuel García-Carpintero, Emiliano Boccardi, Ludovic Soutif, David Suárez-Rivero and Nick Zangwill for more recent discussion and suggestions. Research for this paper was supported by CNPq (Brazil).
2. This is by no means trivial; consider the possibility that we might be talking of two distinct persons called 'Socrates', the philosopher and the soccer player. We may also be talking of the the same person (e. g., the philosopher) but under two completely different registers, one as the character described in Plato's dialogues, and another as the real historic person that lived in Athens. Abbreviations also may, I think, give rise to doubts as to the name identity. E.g., *The I.R.S.= The Internal Revenue Service* has the form *a=a* or *a=b*?
3. Actually one could draw a different conclusion, more similar to Kripke's, i. e., that there is a puzzle (or at least the possibility of a puzzle) quite independently of the particular one raised by Frege. For, of course, it remains open the possibility that one might be puzzled by the fact that 'Paderewski is Paderewski' is surprising in some circumstances, but not in others.
4. I do not mean that no theory at all is involved here, since we are appealing to some theoretical aspects of demonstratives. But the theory is neutral in the sense that the puzzle can be generated without assuming the sense/reference distinction.

References

Frege, G. (1892). Über Sinn und Bedeutung *(SuB)*. *Zeitschrift für Philosophie und philosophische Kritik*, 100/1, S. 25-50. Translated and reprinted in Beaney, M. (Ed.) (1997) *The Frege Reader*. Oxford: Basil Blackweell.

_____. (1918-9). Der Gedanke. *Beiträge zur Philosophie des deutschen Idealismus 2*, 58-77. Translated and reprinted in Beaney, M. (Ed.) (1997) *The Frege Reader*. Oxford: Basil Blackwell.

Glezakos, S. (2009). Can Frege Pose Frege's Puzzle? In Almog, J. and Leonardi, P. (Eds.) (2009) *The Philosophy of David Kaplan*. Oxford: Oxford University Press, 202-207.

Kaplan, D. (1990). Words. *The Aristotelian Society Supplementary*, 64.

Perry, J. (2009). Frege and Hume on Identity. *Philosophical Studies*. 146 (3), 413-23.

(*) **Marco Ruffino** (Department of Philosophy, UNICAMP-Brazil) received a Masters Degree in Logic and Philosophy of

Science from the State University of Campinas (UNICAMP, Brazil), and later a *Ph D* in Philosophy from the University of California, Los Angeles (UCLA) with a dissertation on Frege. His fields of interest include philosophy of language, philosophy of logic, metaphysics and history of analytic philosohy. He is currently professor of philosophy at the UNICAMP.

Received: Monday, October 20, 2014.
Approved: Thursday, October 23, 2014.