

Journal of Philosophy, Inc.

The First Person: Error through Misidentification, the Split between Speaker's and Semantic Reference, and the Real Guarantee

Author(s): Annalisa Coliva

Source: *The Journal of Philosophy*, Vol. 100, No. 8 (Aug., 2003), pp. 416-431

Published by: [Journal of Philosophy, Inc.](#)

Stable URL: <http://www.jstor.org/stable/3655580>

Accessed: 04/09/2010 10:05

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=jphil>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Journal of Philosophy, Inc. is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Philosophy*.

<http://www.jstor.org>

COMMENTS AND CRITICISM

THE FIRST PERSON: ERROR THROUGH MISIDENTIFICATION,
THE SPLIT BETWEEN SPEAKER'S AND SEMANTIC
REFERENCE, AND THE REAL GUARANTEE*

In this article, I wish to contrast three different phenomena concerning the first person, which, so far, have not been sufficiently distinguished in the philosophical literature. A lack of clarity about their distinction has led philosophers such as Elisabeth Anscombe,¹ and, more recently, Carol Rovane² and Andrea Christofidou³ to claim that all, or at least most uses of 'I' and its cognates are *immune to error through misidentification*. I wish to show how such a conclusion is mistaken and possibly brought about, at least in part, by the conflation between error through misidentification and what I will call the *split between speaker's reference and semantic reference*. Further reflection on the latter phenomenon, in turn, will show that while not all uses of 'I' are immune to that split, it is true that they all have a further much more fundamental and important guarantee, so far I believe unremarked, which I will call the *real guarantee*. It is this real guarantee, I suggest, that Anscombe, Rovane, and Christofidou are in fact trying to account for under the misapprehension that it is what others, such as Sydney Shoemaker, Gareth Evans, Crispin Wright, and John McDowell, have called immunity to error through misidentification.⁴

I. ROVANE'S AND CHRISTOFIDOU'S ACCOUNTS
OF ERROR THROUGH MISIDENTIFICATION

As is familiar, the idea of immunity to error through misidentification can be traced back to Ludwig Wittgenstein's distinction, in *The Blue*

* I would like to thank the Fulbright program and the Fonds National Suisse EFRIFO406101443 for their support.

¹ "The First Person," in Samuel Guttenplan, ed., *Mind and Language: Wolfson College Lectures 1974* (New York: Oxford, 1975), pp. 45–64; reprinted in Palle Yourgrau, ed., *Demonstratives* (New York: Oxford, 1990), pp. 135–53 (page reference to the latter).

² "The Epistemology of First-person Reference," this JOURNAL, LXXXIV, 3 (March 1987): 147–67; "Self-reference: The Radicalization of Locke," this JOURNAL, XC, 2 (February 1993): 73–97—here see p. 90; and *The Bounds of Agency* (Princeton: University Press, 1998).

³ "First Person: The Demand for Identification-free Self-reference," this JOURNAL, XCII, 4 (April 1995): 223–34.

⁴ Shoemaker, "Self-reference and Self-awareness," this JOURNAL, LXV, 19 (October 3, 1968): 555–67; Evans, *The Varieties of Reference* (New York: Oxford, 1982); Wright, "Self-Knowledge: The Wittgensteinian Legacy," in Wright, Barry C. Smith, and Cynthia McDonald, eds., *Knowing Our Own Minds* (New York: Oxford, 1998), see pp. 13–45; McDowell, "Reductionism and the First Person," in Jonathan Dancy, ed.,

Book, between what he called uses of 'I' (or 'me' and 'my') *as subject* and its uses *as object*.⁵ Examples of the former kind of use are "I have a toothache," "I see so-and-so"; examples of the latter, by contrast, are "My arm is broken," "I have a bump on my forehead." The difference between the two kinds of use, according to Wittgenstein is this: in the latter case I could utter, "My arm is broken," say after a car accident, on the basis of feeling pain in my arm and of seeing a broken arm at my side. Yet, that self-ascription could be based on having mistaken the broken arm at my side for mine, when, in fact, it is someone else's.⁶ In such an event, there would be someone's arm that is broken but I would be mistaken in thinking that it is mine. By contrast, in the former case, there seems to be no provision for this kind of error. I may be wrong about what I am seeing, for instance, but there is no issue but that it is I who am doing the seeing.⁷

Here, however, is how Rovane characterizes error through misidentification:

Let me describe, therefore, a situation in which we *might* [italics added] be prepared to grant that a use of the first person is accompanied by error through misidentification. Suppose I am facing a mirror and I believe that I see my own reflection when I really see someone else's. On the basis of what I see reflected in the mirror I say, "There's an incredibly tasteless painting hanging on the wall directly behind me." Because I believe that I am the person reflected in the mirror, *I take 'me' to refer to the person reflected in the mirror* [italics added]. Thus there is a sense in which I take someone who is not myself, namely, the person reflected in the mirror, to have made my utterance and, hence, to be the referent of that use of 'me'. *As is usual with cases of error through misidentification, we can interpret the utterance so as to make it come out true or false, depending on what referent we prefer to assign to the offending expression* [italics added].... We are inclined to say that I must have intended to

Reading Parfit (Malden, MA: Blackwell, 1997), pp. 230–50. See also Christopher Peacocke, *Sense and Content* (New York: Oxford, 1983); and J. Pryor, "Immunity to Error through Misidentification," *Philosophical Topics*, xxvi, 1 (1998): 271–304.

⁵ Cf. *The Blue and the Brown Books* (Oxford: Blackwell, 1958), see pp. 66–67.

⁶ In the course of this article, I will mostly talk about self-ascriptions that are liable to error through misidentification, which may or may not be linguistically expressed. I will also use the following conventions: single or double inverted commas for words and sentences, respectively, and square brackets for concepts and propositional contents.

⁷ Nowadays there is a widespread agreement on the fact that immunity to error through misidentification is *not* a phenomenon that is characteristic only of self-ascriptions and that, for instance, uses of demonstratives can also be immune to that form of error. There is also widespread consensus about the fact that, contrary to what seems likely to have been Wittgenstein's position on the matter, the distinction between error and immunity to error through misidentification, relative to the first person, cuts across the physical and the psychological domains.

refer to myself when I used the expression 'me' and that I thought that *in referring to myself I was also referring to the person reflected in the mirror*. Hence it is more natural to interpret my utterance as false of myself rather than as true about the person reflected in the mirror. Nevertheless, *the alternative interpretation might be warranted in certain circumstances, and, when it is, we have a use of the first person that incorporates error through misidentification* [italics added]. We must grant, then that, although uses of 'I' can be subject to referential error through misidentification, it is extremely rare that they are. There is always the strongest presumption that a speaker intends to speak of *herself* in using 'I'.⁸

I have quoted at length because the passage clearly indicates that, on the basis of the example she discusses, Rovane thinks the following:

- (1) that if error through misidentification relative to the first person occurs, then 'I' and its cognates are not used to refer to the subject but to the person who is somehow (in this example, visually) presented to her;
- (2) that—for reasons we will see in a moment—in the case of the use of the first person, it is merely *extremely rare* that it is used to refer to someone else. As a consequence, Rovane holds
- (3) that *most* uses of the first person are immune to error through misidentification.

Before addressing assumption (1), which is the leading thought, let me briefly present and discuss Rovane's various defenses of (2).

Rovane thinks that reference to objects, either in speech or in thought, is always mediated by some belief(s) about the referents, which would determine which object one is speaking or thinking about. Her account of which belief(s) should be in place in order to secure self-reference, however, varies across her writings. On the one hand, it is held that "speakers generally have enough true beliefs about themselves in virtue of which they are quite clear about their identities even though they have some false beliefs about themselves as well."⁹ Hence, it is because it is very seldom the case that one has *massively* false beliefs about oneself that there is always the strongest presumption that in using the first person one is thinking or talking about oneself. On the other hand, since, on this view, it could hardly be explained how amnesiac and seriously deluded subjects could manage to refer to themselves by using 'I', Rovane presents two other identity beliefs to mediate first-person reference, namely [I = the

⁸ "The Epistemology of First-person Reference," pp. 153–54. Rovane introduces error through misidentification essentially in the same way in her more recent work—cf. *The Bounds of Agency*, pp. 220–21.

⁹ "The Epistemology of First-person Reference," p. 154.

person who is speaking now] and [I = the thinker of this thought].¹⁰ But these beliefs seem to be a priori true, for, whenever I speak, it is true that I am the person who is speaking now. Moreover, if it is a priori true that I can only be introspectively aware of my own thoughts,¹¹ then the identity between *I* and *the thinker of this thought* would hold a priori. Yet, on both these views, assuming further that (1) is correct, *all* uses of the first person would be immune to error through misidentification, contrary to what Rovane maintains—that error through misidentification *can*, exceptionally, afflict uses of the first-person pronoun and its cognates.

This is in effect Christofidou's position, although in her article she reacts to Rovane's account on the grounds that the *reliability* of [I = the thinker of this thought] would never secure *immunity* to error through misidentification, but only, at most, that one's self-ascriptions—which remain potentially open to that kind of error—are in fact not affected by it (*op. cit.*, p. 226). Here are two relevant passages:

Even if my memories and perceptions deceive me, and it is not *I* who climbed the steps of St. Paul's but a friend, then I misidentify, not myself, but the person who climbed the steps; my use of 'I' refers correctly to me, and refers to me directly or unmediatedly. In discovering that I did not after all climb the steps of St. Paul's, then I should say 'so I did not climb them', and then I should try to explain how I could have fallen into error. But the error has nothing to do with the self-reference of 'I'. If there is a problem with propositions such as 'I climbed the steps of St. Paul's' the mistake is in the predication component, *not in the identification component*. I have *misascribed* to myself a predicate, but the reference or identity of 'I' is unshaken: 'I' is immune to error through misidentification, whatever the predicate might be (*op. cit.*, p. 227; italics added).

[D]iscussions in this area have been vitiated by the failure to distinguish between two types of immunity: immunity to error through misidentification (which covers *all* uses of 'I' whether these involve corporeal predicates, or mental predicates, or both) and what I call immunity to error through *misascription* (which covers only certain mental predicates) (*op. cit.*, footnote 7).

Here again, however, the startling claim

(3') that *all* uses of 'I' are immune to error through misidentification is reached by holding

¹⁰ The first proposal is made in "The Epistemology of First-person Reference," and in *The Bounds of Agency*, pp. 221–22; the second in "Self-reference."

¹¹ I defend this claim in my "Thought Insertion and Immunity to Error through Misidentification," *Philosophy, Psychology and Psychiatry*, ix, 1 (2002): 27–34.

- (1) that if error through misidentification occurs, then 'I' would not refer to the subject but to someone else and
 (2') that this is impossible, because 'I' refers directly and unmediatedly to the subject.¹²

Now, claim (2') allows of two interpretations. It can be taken to mean that 'I' is an *automatic indexical*, which invariably refers to the producer of its tokens.¹³ Hence, all tokens of 'I' refer to the person who produced them. Or else, it can be taken to mean that since linguistic self-reference, in a framework like the one elaborated by Evans in *The Varieties of Reference* (*op. cit.*), for instance, is taken to be dependent on the ability to have I-thoughts and these, in turn are grounded on *identification-free* ways of having knowledge about oneself, then *all* self-reference by means of the first person is direct and unmediated.¹⁴ I surmise that Christofidou has this second reading of (2') in mind. Yet, it is striking—and indeed it should have struck Christofidou—that Evans for one *never* claimed that *all* uses of 'I' were immune to error through misidentification, while, for the reasons just given, he indeed thought that linguistic *self-reference* was unmediated.

Before turning to a discussion of assumption (1), let me point out an important difference between Christofidou's treatment of the relevant class of self-ascriptions and Rovane's. Christofidou's view—contrary to Rovane's—allows them to be false, because, on her view, 'I' would always refer to the subject who, however, may not instantiate the property expressed by the predicate. Yet she thinks that that would be just a common mistake of *misascription*—as opposed to a mistake of misidentification—relevantly similar to the case in which, for instance, being deluded about the color of the shirt I am wearing, I falsely say, "I am wearing a white shirt."

Suppose, however, that my claim "I am wearing a white shirt" was

¹² In fact, the particular example given by Christofidou is problematic. For on some interpretations (cf. McDowell) it would turn out to be immune to error through misidentification, while on others (cf. Pryor), it would not. It all depends on how we choose to interpret cases of quasi-memories and on what account we give of Shoemaker's distinction between *de facto* and logical immunity to error through misidentification, if we accept it at all. This, however, is not an issue that I can pursue within the scope of this article.

¹³ Cf. John Perry, "Indexicals and Demonstratives," in Bob Hale and Wright, *A Companion to the Philosophy of Language* (Malden, MA: Blackwell, 1997), pp. 595–98 for the use of the label "automatic indexical" and its discussion.

¹⁴ Knowledge of the truth of a thought such as [I am *F*] is identification-free, according to Evans, if it is not based on knowledge of the truth of a pair of thoughts such as [*a* is *F*] and [I = *a*], but is just based on receiving information directly about oneself, either through perception and proprioception, or through introspective awareness of one's own mental properties, or by having stored information about oneself in memory—cf. Evans, chapter 7.

made on the basis of seeing someone reflected in a mirror wearing a white shirt and of wrongly taking that person to be me. Suppose, further, that, by chance, it was really the case that I was wearing a white shirt. Now, on Christofidou's account of the matter, since we cannot say that my use of 'I' is affected by error through misidentification, we are obliged to say that my claim is perfectly in order, for, in this case, there is no misascription of a predicate, after all. But this seems to be highly counterintuitive. Again, this counterintuitive view would depend on having excluded the possibility that uses of 'I' can be affected by error through misidentification, on the ground that if they were, they would not refer to the subject.

To sum up: despite their differences, Rovane's and Christofidou's accounts of error through misidentification and of the corresponding kind of immunity are based on a structurally similar piece of reasoning, which assumes that if error through misidentification occurs, then 'I' and its cognates would be used to refer to someone other than oneself. Recoiling from this idea, they try to qualify the conclusion by, respectively: attempting a semantic explanation (Rovane), appealing to the extraordinary reliability (which is in fact the a priori truth) of the identity belief that allegedly mediates first-person reference, namely [I = the thinker of this thought]; and by denying (Christofidou) that any belief mediates first-person reference at all.

Both their accounts seem problematic, however, for a striking consequence of their views is that, for *conceptual* reasons, error through misidentification is never or hardly ever possible. This runs contrary to what has been variously maintained in the rest of the literature on the topic,¹⁵ and to some solid intuitions. For, after all, it seems quite intuitive to say that, in Rovane's own example, for instance, the claim "There is an incredibly tasteless painting hanging on the wall behind me" would be affected by error through misidentification, relative to the first person, if the person reflected in the mirror were in fact someone else. In such an event, my claim would not be false because the person reflected in the mirror does not instantiate the property ascribed to her by the predicate. It would be false *because I am not that person*.

II. ERROR THROUGH MISIDENTIFICATION: THE CORRECT ACCOUNT

Let us now turn to assumption (1), on which both Rovane's and Christofidou's accounts of self-ascriptions that are immune to error through misidentification ultimately rest: the assumption that if error through misidentification occurred (or could occur), then one would

¹⁵ Cf. footnote 4.

not be saying something about oneself, but about the person one has mistakenly taken to be oneself, so that 'I' (or its cognates) would refer not to its utterer, but to someone else.

In order to appraise this assumption, let me propose a template. Consider Rovane's example again. Here is how the subject's mistake can be accounted for. The self-ascription "There is an incredibly tasteless painting hanging on the wall behind me" can be seen as rationally *grounded* on the following pair of beliefs: [There is an incredibly tasteless painting hanging on the wall behind the person reflected in the mirror], and [I am the person reflected in the mirror]. Call these beliefs the *predication* and the *identification component*, respectively. These beliefs rationally sustain the subject's final self-ascription in the following sense: while she may not consciously entertain them in order to move to her self-ascription, still we would expect her to appeal to them in order to justify it, if challenged, and to withdraw from it if one of the two sustaining beliefs turned out to be false.¹⁶

In Rovane's view, as we have seen, the resulting self-ascription may be regarded as true or false depending on whether we take the use of 'I'/'me' to refer either to the person actually reflected in the mirror, or to the subject. Rovane also insists, however, that, although it is more natural to take the final self-ascription as saying something false of the subject, "the alternative interpretation might be warranted in certain circumstances, and, when it is, we have a use of the first person that incorporates error through misidentification."¹⁷ This is odd, however. For, if this claim is true, what does the error consist in? There seem to be three possibilities.

According to a first reading, we might take Rovane to be saying the following: the final self-ascription is affected by error through misidentification because it is grounded on a false identification component; yet it is true because 'I' should be taken to refer to the person who is actually reflected in the mirror. This, however, would require that the first person be used ambiguously across the inference. For the identification component is wrong only if the occurrence of the first person in it refers to the person standing in front of the mirror, rather than to the person she sees reflected in the mirror. The cost of this interpretation is accordingly that the resulting self-ascription would be unwarranted, because it would be based on an equivocation between the occurrence of the first person in the premise and its occurrence in the conclusion. More explicitly, on this first reading

¹⁶ For a discussion of this issue, cf. Peacocke, pp. 143–44.

¹⁷ "The Epistemology of First-person Reference," pp. 153–54.

we can represent the inferential grounds, which rationally sustain the self-ascription, as follows:¹⁸

- (a)(I) A is F
- (II) $A = I_B$
- (III) I_A am F

where, unbeknownst to the subject, $A \neq I_B$. However, (a) is not valid, since the use of the first person in the second premise and in the conclusion would refer to two different people. So, on this reading, any case of error through misidentification would be a case of *faulty reasoning leading to a true belief*.¹⁹

According to the second possible reading, we interpret the use of the first person across the inference as uniformly referring to the person visually presented to the subject. As a consequence, however, we would have that the identification component on which the final self-ascription is based is, surprisingly, correct. For if the first person refers to the person who is actually reflected in the mirror, then surely she is identical to herself. Yet, on this reading, we would have the startling conclusion that error through misidentification occurs if a self-ascription is based on a *true* identification component!²⁰ The form of inference that would rationally sustain the self-ascription would be as follows:

- (b)(I) A is F
- (II) $A = I_A$
- (III) I_A am F

(b) would thus be valid and (III) true. But, then, where is the error?

¹⁸ Where ‘ A ’ refers to the person actually reflected in the mirror, and ‘ B ’ refers to the subject who is making the self-ascription, and the subscripts indicate who is actually referred to by an occurrence of the first person.

¹⁹ In conversation, Rovane has distanced herself from this reading and has proposed the following:

- (I) A is F
- (II) $A =$ The person I take to be me
- (III) The person I take to be me = I_B
- (IV) I_B am F

This inference is valid alright, but the reader will be able to verify that it would in fact be a special case of (c) below and that the conclusion would be *false*, contrary to what Rovane claims should happen when error through misidentification occurs.

²⁰ The same kind of arguments would apply against Christofidou. For, if error through misidentification *could* occur, then the final self-ascription would not be about myself, but about the person reflected in the mirror. But this could only be the case if either the self-ascription is based on an equivocation between the occurrence of the first person in the premise and in the conclusion, or else, if it is based on a correct identification component, which is absurd.

There is only one last possibility: to adopt a uniform reading of the first person across the inference, yet, this time, as referring to the subject who is making the self-ascription. The result would be that the final self-ascription is false and it is so because it is based on the false identification component [$I =$ the person reflected in the mirror]. The form of inference, which would rationally sustain the final self-ascription, can then be represented as follows:

- (c) (I) A is F
 (II) $A = I_B$
 (III) I_B am F

(c) is valid. Yet, unbeknownst to the subject, $A \neq I_B$. Hence, the conclusion is false.

This, I submit, is the right account of the matter. Yet, to require that, in order to account for error through misidentification, the first person be interpreted across the inference as uniformly referring to oneself, is actually *inconsistent* with both Rovane's and Christofidou's assumption (1) according to which, if error through misidentification occurs, then the use(s) of the first person would refer to someone else.

A satisfactory account of error through misidentification relative to the first person should then respect the following features:

- (1) error through misidentification can occur, although it may be contingently rare;
- (2) when it does, it may issue in a false belief, correctly expressed;
- (3) yet that belief is rationally grounded and this requires
- (4) continuity of reference to oneself throughout the belief and its grounds.

Returning to Rovane's example, this means that "There is an incredibly tasteless painting hanging on the wall behind me" is affected by error through misidentification precisely because it is a statement about *me*—and not about the person reflected in the mirror whoever she is. Therefore, that self-ascription is false because it is not true of *me* that I instantiate the property expressed by the predicate and I make it because I wrongly take the person reflected in the mirror to be me.

III. THE DIAGNOSIS I: THE SPLIT BETWEEN SPEAKER'S AND SEMANTIC REFERENCE

The question we should now ask is why one may be tempted to hold Rovane and Christofidou's assumption (1). My suggestion is that this may be at least partly due to a conflation between error through misidentification and the second phenomenon I am here concerned

with: the (im)possibility of a split between speaker's and semantic reference.

Anscombe, in her famous paper on the first person, argues that 'I' has certain guarantees, which other singular expressions do not have. In particular, she claims that uses of 'I' are guaranteed of having a referent, because the very utterance of a token of 'I' presupposes and, therefore, guarantees its existence.²¹ She also claims that any given use of 'I' has a further guarantee as well: that the person that the user takes to be its referent is its referent. She phrases her point as follows:

Guaranteed reference for...['I']...would entail a guarantee, not just that there is such a thing as X, but also that what I take to be X is X (*op. cit.*, pp. 145–46).

Anscombe calls this latter guarantee "Immunity to error through misidentification." She goes on to characterize it as follows:

An 'I'-user cannot take the wrong object to be the object he means by 'I'. (The bishop may take the lady's knee for his, but could he take the lady herself to be himself?) (*ibid.*).

Anscombe's thought could appear to be this, that for any given use of 'I' to be comprehending, it must be such that one uses it intending to refer to oneself. Since, according to Anscombe, one cannot take someone else to be oneself, any use of 'I' is guaranteed of reference to its proper semantic referent, namely oneself.

Contrast 'I', so viewed, with a proper name like 'Dummett'. For a token use of the latter to be comprehending it must be made intending to refer to Dummett, but since I can take the wrong person to be Dummett, my use of that name is not guaranteed of reference to its actual semantic referent. The idea is that one can use a proper name comprehendingly, that is, intending to refer to its semantic referent, while not being able to answer the question, "Which object is *a*?" for instance, by saying (correctly), "That (perceptually presented) person is *a*." This, in turn, opens up the possibility that one may use 'Dummett' intending to refer to Dummett, but, in effect, not knowing which person is Dummett, take the wrong person to be him and use that name to refer to that person. In short: failure of knowing which person is the semantic referent of a singular expression allows one to use that expression to refer to someone other than its semantic

²¹ Evans notoriously contests this guarantee, but this is not the place to discuss this issue.

referent, if the latter is mistakenly identified with the wrong object. Call this phenomenon *the split between speaker's and semantic reference*. Such a split was made familiar by Keith Donnellan,²² although his classic discussion restricted it to the use of definite descriptions.

Anscombe's use of "immunity to error through misidentification" thus applies to something we might better call "impossibility of a split between speaker's and semantic reference." For it consists in the (putative) guarantee that if one's use of 'I' is competent, one must intend to use it to refer to oneself and that, in effect, the object one takes to be oneself is indeed oneself. This guarantees that the proper semantic referent of 'I'—namely, oneself—and the object that one intends 'I' to refer to, on a given occasion—that is, the "speaker's referent" of 'I'—be one and the same.

Now, a conflation between error through misidentification and a split between speaker's and semantic reference may well provide at least part of the motivation for the claim that if error through misidentification could arise, then one's use of 'I' would not refer to oneself but to the person who is wrongly taken by the subject to be oneself. Recall, however, that Anscombe (as well as Rovane and Christofidou) holds that this may never (or almost never) occur, in the case of the use of 'I' *on conceptual grounds*.

On reflection, however, it is obviously not true that all uses of 'I' are immune to a split between speaker's and semantic reference. For, while intending to use 'I' to refer to myself, I can wrongly take someone else for me, on a given occasion, and, hence, use 'I' to speaker-refer to that person. The usual kinds of example suffice to note this point. Consider, for instance, the case of a person who uses 'me' to refer to someone, different from herself, she sees reflected in the mirror. If there is a sense in which she can take that person to be herself, then there is a sense in which her use of 'I' can refer to that person. I think there is such a sense. For suppose that someone had just said, "There are no tasteless paintings in this room," and that, seeing the painting seemingly hanging behind her, the speaker feels in a position to contradict her interlocutor, by pointing out to him that, after all, there is one incredibly tasteless painting hanging on a wall of the room. In that case, if he then pointed out to her that the

²² "Reference and Definite Descriptions," *Philosophical Review*, LXXV, 3 (1966): 281–304. Although Rovane is skeptical about the distinction between speaker's and semantic reference, she introduces error through misidentification in connection with Donnellan-like examples, which, in my view are best accounted for as cases of such a split—cf. "The Epistemology of First-person Reference," p. 152; and *The Bounds of Agency*, p. 220.

person behind whom there is a tasteless painting is not her, she would reply, "Well, there is an incredibly tasteless painting hanging on the wall behind *that* person reflected in the mirror." What her reply would show is that her original referential *intentions* in using 'I' were *primarily* directed towards that person, whom she mistakenly took to be herself. So, it is not true that all uses of 'I' and its cognates are immune to a split between speaker's and semantic reference. Once again, there does not seem to be any conceptual barrier to the possibility that such a split, however contingently rare it may be, can afflict the use of 'I'. Hence, to sum up:

- (i') a split between speaker's and semantic reference can occur in the case of the use of 'I', although it may be contingently rare;
- (ii') when it does, it may result in a true statement which
- (iii') is based on an ungrounded inference, and
- (iv') this, in fact, requires that the occurrence of the first person in the conclusion be used (in the kind of example I am considering here) to speaker-refer to an object other than oneself, somehow presented to one.²³

So if what Anscombe, Rovane, and Christofidou are trying to highlight is some kind of sure-fire guarantee for our uses of 'I' and 'I' is immune neither to error through misidentification, nor to the split between speaker's and semantic reference, then we still have to find out what it is.

IV. THE DIAGNOSIS II: THE REAL GUARANTEE

In order to get clear about what Anscombe, Rovane, and Christofidou may be after, I think it is crucial to realize that there is an important difference between the case of a split between speaker's and semantic reference, which may occasionally afflict uses of 'I', and the case where that split afflicts uses of other singular expressions—proper names and definite descriptions in particular.²⁴ The difference is this. In the latter case (usually, although not necessarily)²⁵

²³ Notice how the pattern of inference (a) presented in the previous section would fit with this account of the split between speaker's and semantic reference.

²⁴ Even someone who was skeptical about the possibility of a split between speaker's and semantic reference in the case of 'I' should acknowledge that *if* it could occur, then it would be structurally different from the same kind of split arising for other singular expressions.

²⁵ This, however, does not make any difference for our present purposes because, for the contrast to stand, it is sufficient that this *may* be the explanation of the split, at least in some cases of uses of proper names and definite descriptions, while it can *never* be the explanation of the split in the case of the use of 'I' (provided, as we will see, that 'I' is used comprehendingly).

- (a') the subject does not know which person is the *semantic* referent of the singular term;
- (b') this ignorance explains why she may take another person, who is different from the semantic referent of the singular term, to be the referent of that expression.

By contrast, in the former case, although a split can occur, that does not threaten the *presumption* that

- (a'') the subject indeed knows which person she is.

Hence, ignorance of which person is the *semantic* referent of 'I' cannot be appealed to in order to explain why a subject may take someone else to be the person to whom a given use of 'I' by her refers. So, what explains it?

The correct form of explanation can be illustrated as follows:

- (a*) the subject (to continue with Rovane's example) is ignorant of the identity of the *person visually presented to her*, and
- (b*) she mistakenly identifies *that person* as herself.

The point is simply that one may well be able visually to discriminate a person—*that person*—yet still not be in a position to answer the question, "Which person is that one?" or, equivalently, "Who is that person?" Then, having wrongly identified that person as oneself, one goes on to use 'I' to speaker-refer to that person. Yet, as Christofidou would put it, one has not misidentified oneself, but *that person*. That is to say, although one may use 'I' in order to purport reference to that person because one does not know who that person is and wrongly takes her to be identical to oneself, one is under no misapprehension about *one's own* identity in doing so.

The guarantee possessed by competent uses of 'I' is then, not immunity to error through misidentification, nor to a split between speaker's and semantic reference, but this:

The Real Guarantee (at the level of language): the comprehending use of 'I' guarantees that the speaker knows which person is its *semantic* referent.

The real guarantee holds on purely conceptual grounds. It is a rule for the competent use of 'I' that one must use it to refer to *oneself*. But one cannot so much as be in a position to understand this rule unless one knows which person one is. This, in turn, implies that *qua* competent 'I'-user, the subject knows which person is the semantic referent of 'I'.

Notice, however, that the real guarantee must also operate at the level of thought. Hence, even if one were skeptical about the distinc-

tion between speaker's and semantic reference in the case of 'I', one could still retain the following:

The Real Guarantee (at the level of thought): the possession of the first-person concept guarantees that the subject knows which person that concept is a concept of.

The real guarantee holds at the level of thought because the first-person concept is the concept of *oneself* and one cannot have it unless one knows which person one is. Hence, there is a guarantee that *qua* possessor of the first-person concept, the subject knows which person that concept is a concept of.

My claim then is that the phenomenon which Anscombe and, more recently, Rovane and Christofidou have really been concerned with, is not immunity to error through misidentification, nor immunity to a split between speaker's and semantic reference, but in fact the real guarantee. Indeed, a profitable way of interpreting the debate between Rovane and Christofidou is to see it as concerning how best to characterize one's knowledge of which person one is, which, in its turn, is supposed to ground one's grasp of the first-person concept both in language and in thought: while Rovane thinks of this knowledge as a kind of propositional knowledge of the form "I am *F*," or "I am the *F*," Christofidou thinks of it as some kind of nonpropositional, direct acquaintance with oneself.

Looking at their debate from this perspective can in fact help us to arbitrate it. For, if knowledge of oneself depended on having a set of (mostly) true beliefs about oneself, then it would follow that global amnesiacs would not know which persons they are. Hence, they could neither use 'I' comprehendingly, nor have thoughts about themselves. But, surely, we would like to be able to say that if a global amnesiac said (or thought), "I cannot remember anything about my past," her use of the first person would be perfectly in order. This, in turn, shows that knowing which person one is is *not constitutively* dependent on having such a set of true beliefs about oneself.²⁶

²⁶ Indeed, if this proposal were meant to imply that, had one a set of false beliefs about oneself, then one might in fact be using the first person to think and talk about some other person of whom those beliefs are true, it would be doomed to failure for the following reason. In order for beliefs expressible as "I am *F*," or "I am the *F*" to be false, they should indeed be about *oneself*. Otherwise, if they were about the person of which they are true, they would be true after all! Hence, in using the first person (either in speech or in thought) in entertaining those beliefs, the subject must indeed be referring to, or thinking about herself, and, to that extent, she must know which person she is. What the subject would then be doing, as Christofidou claims, is to misascribe a given property to *herself*, namely, either the property *F*, or the property of being identical to *F*.

Notice, however, that this consequence may in fact be blocked by Rovane, by

Moreover, the other proposals, according to which knowledge of which person one is is either given by the belief expressible as "I am the person who is speaking now," or by the belief expressible as "I am the thinker of this thought," face two immediate problems. First, any suggestion that construes knowledge of oneself in terms of *beliefs about oneself* presupposes a grasp of the first-person concept already, hence knowledge of which person one is. Therefore, it cannot be an explanation of what that knowledge amounts to. Second, such a suggestion would not have the resources to discriminate between the subject and its possible doppelgänger. As a consequence, the very possibility of genuine, singular I-thoughts would be impaired.²⁷

Thus, I take it that Christofidou is right in thinking of knowledge of oneself as essentially unmediated by any belief about oneself and as depending instead on some kind of *acquaintance with oneself*. In fact, Evans, followed by McDowell and later by Peacocke, construed knowledge of oneself as a kind of acquaintance with one's own body, its environment, one's memories and one's own mental states. Although this is not the place to rehearse the details of their respective proposals, the important point about their views is just this: knowledge of which person one is consists in the *dispositional* ability to form thoughts about oneself on the basis of the *information* deriving from one's body, its environment, one's memory, and from one's own mental states. Hence, although the amnesiac would not actually be able to form I-thoughts on the basis of her memories, nor would an anaesthetised subject actually be able to form them on the basis of her proprioceptive feelings, or of the perception of the environment around her, they may have nevertheless retained the relevant dispositions.²⁸ These, in turn, once combined with one's ability to form I-thoughts on the basis of one's own occurrent mental states, would suffice for giving those subjects knowledge of which persons they are. Knowledge of which persons they are would then secure that any (comprehending) use of the first person either in speech or in thought would possess the real guarantee.

holding that, no matter whether one has massively false beliefs about oneself, one would still know which person one is if one had the presumably unimpaired beliefs expressible as "I am the person who is speaking now," and "I am the thinker of this thought."

²⁷ I would like to thank Rovane for generously drawing this latter problem to my attention.

²⁸ As L. O'Brien notices: "One should not conflate a non-activated disposition with the absence of that disposition at all"—"Evans on Self-Identification," *Noûs*, xxix, 2 (1995): 232–47, here see p. 237.

V. CONCLUSIONS

Anscombe, Rovane, and Christofidou have all been after the idea that in our use of the first person, we possess some kind of distinctive and sure-fire guarantee. Yet, on reflection it has become apparent that it cannot be the case that that guarantee is either against error through misidentification, or against the possibility of a split between speaker's and semantic reference, for it is enough for a particular use of 'I' to be affected by the former error that that pronoun is used to give expression to a judgment about oneself which is based on a belief in a wrong identification component. And it is enough for a particular use of 'I' to be affected by the latter kind of mistake that one has the intention primarily to refer to a person somehow demonstratively presented to one, and wrongly taken as oneself.²⁹

The real guarantee possessed by our uses of the first person is different, and more simple. It is the guarantee that we each of us know which person is the semantic referent of our tokenings of 'I' and which person our respective first-person concepts are concepts of. It is this guarantee that Rovane and Christofidou are aiming to account for, with their respective proposals about the nature of knowledge of oneself, under the misapprehension, which can be traced back to Anscombe, that it is what others have called immunity to error through misidentification. But we can set that misapprehension aside. The fundamental issue which their accounts should be seen as addressing is the epistemological nature of self-reference, whether in speech or in thought, and how it underwrites the real guarantee.

ANNALISA COLIVA

University of Bologna

²⁹ Notice that the difference between these two kinds of error is exhausted by the kind of (primary) referential intentions involved. Namely, the (primary) intention of saying something about oneself in the former case, and the (primary) intention of saying something about the person visually presented to one, in the latter case.