

The second half of Doyle's book is devoted to a subtle and often ingenious discussion of Anscombe's 'The First Person', in which she notoriously concludes that 'I' is not a referring expression. Doyle's aim is not really to defend this conclusion, though he does devote a chapter to developing an anti-referentialist account of 'I' in more detail than Anscombe herself does (only then to elaborate some serious problems the view faces). Instead, Doyle's central concern is to argue that no existing referentialist account of 'I' overcomes a particular problem raised by Anscombe—a problem Doyle calls this the 'circularity problem'.

Doyle argues persuasively that we should take Anscombe's claim that 'I' is nonreferring at face value. Yes, she really does mean it when she says 'I' does not refer to anything. But it is worth emphasizing, in trying to get a better grip on her objections to a referentialist treatment of 'I', that Anscombe does not take this view to have some of the implications it might at first glance seem to have. In particular, Anscombe concedes that statements of the form "I am F" not only (1) have the same *grammatical* form as "a is F", where 'a' is a referring expression (like a name or demonstrative), but also (2) that the two types of sentence are *semantically* alike, in that "I am F" is true (or at least "verified") iff the speaker is F.

But Anscombe denies that conditions (1) and (2) are sufficient for 'I' to be a referring expression designating the speaker. The syntactic criterion (1) is insufficient because 'it' in 'It is raining' does not refer, despite its grammatical resemblance to 'It is ringing', in which 'it' refers to a particular thing that is said to be ringing. The semantic criterion (2) is insufficient because the semantics of individual terms cannot be read off of the semantics of the sentence. As Anscombe puts it, "Of course we must accept the rule 'If X asserts something with "I" as subject, his assertion will be true if and only if what he asserts is true of X.' But ... the truth-condition of the whole sentence does not determine the meaning of the items within the sentence." To illustrate, Anscombe suggests without much elaboration that one might add a terminal -O to a predicate to make a sentence that is verified if the predicate applies to the speaker, without the terminal -O being a referring expression.

I will mostly leave it open whether these concessions, especially the second, do anything to limit the interest of Anscombe's anti-referentialist thesis. I am not sure they do, but it's a reasonable concern.

So, if (1) and (2) are insufficient for 'I' to refer, then what else does it take? Anscombe says it is this: The 'I' user must *intend* to use it to refer to something. For what it's worth, this strikes me as plausible but not entirely obvious. But it is a standing assumption throughout Anscombe's discussion (and Doyle's) and I'll grant it for sake of argument.

In fact, Anscombe assumes something a bit stronger than this. For 'I' to refer, it is not enough for the speaker to intend to refer to *something*. Rather, the speaker must intend to use 'I' as a device of *self*-reference. And this is where the circularity problem comes in. Here is how Anscombe puts it:

Nor can we explain the matter, as we might suppose, by saying "I" is the word each one uses when he knowingly and intentionally speaks of himself.' For did not Smith knowingly and intentionally speak of Smith? Was not the person he intended to speak of—Smith? and so *was* not the person he intended to speak of—himself?

It may be said: 'Not in the relevant sense. We all know you can't substitute every designation of the object he intended to speak of and keep the statement about his intention true.' But that is not the answer unless the reflexive pronoun itself is a sufficient indication of the way the object is specified. And that is something the ordinary reflexive pronoun cannot be. Consider: 'Smith realizes (fails to realize) the identity of an object he calls "Smith" with himself' If the reflexive pronoun there is the ordinary one, then it specifies for us who frame or hear the sentence, an object whose identity with the object he calls 'Smith' Smith does or doesn't realize: namely the object designated by our subject word 'Smith'. But that does not tell us what identity Smith himself realizes (or fails to realize). For, as Frege held, there is no path back from reference to sense; any object has many ways of being specified, and in this case, through the peculiarity of the construction, we have succeeded in specifying an object (by means of the subject of our sentence) without specifying any conception under which *Smith's* mind is supposed to latch on to it. For we don't want to say 'Smith does not realize the identity of Smith with Smith'.

We only have to admit a failure of specification of the intended identity, if we persist in treating the reflexive in 'He doesn't realize the identity with himself' as the ordinary reflexive. In practice we have no difficulty at all. We know what we mean Smith doesn't realize. It is: 'I am Smith.' But if that is how we understand that reflexive, it is not the ordinary one. It is a special one which can be explained only in terms of the first person.

How central is the circularity problem for Anscombe's overall argument? For all its insight and interest, Anscombe's paper is difficult to interpret even on this simple dialectical question. I think I answer it differently than Doyle does, though perhaps it is mostly a difference of emphasis. Doyle's discussion primarily concerns this problem, which Anscombe lays out early in her paper. I think Doyle might read this passage as giving a general outline of the problem she will develop throughout the rest of the paper. But as I read her, Anscombe intends this as an objection to a *particular* referentialist view. The failure of this first-pass referentialism motivates the central questions that more sophisticated referentialist views try to answer, and Anscombe spends the rest of the paper objecting to those views on *distinct* grounds.

The initial referentialist view says something like this:

Simple Referentialism: 'I' is a term a speaker intentionally uses to refer to himself, *and this fully explains the meaning of 'I'.*

Anscombe goes on to deny even the first clause of the sentence. But the circularity problem is a problem only for the additional italicized clause, as I read it.

To see the problem, we must note that this view obviously fails if 'himself' is understood as an ordinary reflexive. On this reading, all that is required for first-person reference is for the speaker to intend to refer to a person who is in fact the speaker himself. But this seems obviously insufficient for using 'I' with its usual meaning. For example, suppose a speaker intends to use 'I' as a nickname for whoever is spilling the sugar, and unbeknownst to him he is the one spilling the sugar. This speaker seems not to be using 'I' with its usual meaning, though he does intend to refer to a person who is in fact himself.

The second reading takes ‘himself’ to be an indirect reflexive. It’s not enough for a speaker to intend to refer to someone who is in fact the speaker himself. The speaker must intend to refer to himself, where ‘himself’ is an indirect reflexive. But what does this amount to? Perhaps it could be suggested that the speaker must intend to refer to someone who he *knows* to be himself. But for me to know of the individual DB that he is myself, at least in the relevant sense, is just to know that I am DB. More generally, intending to refer to myself, in the relevant sense, seems to be an essentially first-personal intention. And this, Anscombe suggests—at least on Doyle’s reading—circularly explains first-person reference in terms that presuppose first-person reference.

What exactly is the problem supposed to be? And what view is it a problem for? Doyle takes it to be a problem of explanatory circularity. The referentialist wants to explain first-person reference in terms of the speaker’s intentions for ‘I’, but you can’t spell out the relevant intentions in a non-circular way. And Doyle thinks this is a very general problem, and he argues with great ingenuity that a number of attempts to overcome it are unsuccessful. I think on his reading, it is the crux of Anscombe’s argument against referentialism, and in favor of her alternative view that ‘I’ is nonreferring.

The passage Doyle quotes lends some support to his reading. The quoted passage ends with an apparent statement of its upshot: “If that is right, the explanation of the word ‘I’ as “the word which each of us uses to speak of himself” is hardly an explanation!” Anscombe then continues “At least, it is no explanation if that reflexive has in turn to be explained in terms of ‘I’; and if it is the ordinary reflexive, we are back to square one.” All of this sure makes it sound like Anscombe is alleging a problem of circularity against at least some referentialist views. But here is how she continues: “We seem to need a sense to be specified for the quasi-name ‘I’. To repeat the Frege point: we haven’t got this sense just by being told which object a man will be speaking of, or whether he knows it or not, when he says ‘I’. ... [I]f I expresses a way its object is reached by him, what Frege called an ‘Art des Gegebenseins’ [*mode of presentation*], we want to know what that way is and how it comes about that the only object reached in that way by anyone is identical with himself.”

As I read her, Anscombe is using the so-called “circularity problem” merely to argue that referentialists are committed to there being a distinctive first-person mode of presentation associated with ‘I’, which determines its reference. Such a view can still allow ‘I’ to be an expression speakers intentionally use to refer to themselves, but it will have to drop simple referentialism’s claim that this is all there is to the meaning of ‘I’. The meaning of ‘I’ is instead given by the relevant mode of presentation, or conception, that each speaker has of himself, and expresses using ‘I’.

As I read it, the rest of Anscombe’s paper argues against *this* view, and her objections to it have no apparent connection to the circularity problem. Instead, the objections have to do with immunity to error through misidentification, and with the apparent conceivability of ‘I am not X’ for any object X—phenomena which Anscombe thinks cannot be explained if ‘I’ is a referring term whose reference is fixed via a mode of presentation in the usual way.

But I’m not going to belabor this interpretive point here. Instead, I want to suggest that Anscombe is *right* not to regard this as a deep problem of circularity, assuming she didn’t so regard it. Let’s grant Anscombe what she claims in setting up the “problem of circularity”.

That is, grant that the referentialist must claim that what secures self-reference is the speaker's intention for her utterance of 'I' to refer to herself, and the relevant intention must itself be distinctively first-personal. This passes the explanatory buck. We want to understand first-person reference in speech, and it explains it in terms of first-person reference in *thought* (specifically, in intentions). But that's not circular! It's very plausible that this is how the explanation should go. Maybe the view would be circular if it tried to explain first-person reference in thought in terms of a further intention one has to have self-referential thoughts. But I think very few would be tempted to explain the contents of our thoughts in terms of further thoughts/intentions we have about what those thoughts will refer to, and Anscombe never appears to consider such a view. So I'm not inclined to think that this referentialist view is explanatorily circular.

Now an extended defense of such a view would require responding to what I see as Anscombe's main objections to it, involving immunity to error through misidentification and the like. I'm not going to do that, since I am not convinced of the view myself, and in any case it would go too far beyond Doyle's discussion. But let me say a little bit about it, before getting back to some points that are more squarely in Doyle's sights.

As we have already noted, Anscombe grants that when I say 'I am F', what I say is true iff David Barnett is F. But why of all people is DB the one whose F-ness matters for the truth of 'I am F' as spoken by me? Anscombe doesn't take this to be some brute fact. Instead, she says:

There is a real question: with what object is my consciousness of action, posture, and movement [connected?], and are my intentions connected in such a fashion that *that* object must be standing up if I have the thought that I am standing up and my thought is true? And there is an answer to that: it is this object here.

'I am this thing here' is, then, a real proposition, but not a proposition of identity. It means: this thing here is the thing, the person ... of whose action *this* idea of an action is an idea, of whose movements *these* ideas of movement are ideas, of whose posture *this* idea of posture is the idea. And also of which *these* intended actions, if carried out, will be the actions.

So, I say something true by 'I am DB', and by 'I am standing' I say something that is true whenever DB is standing. I say something false by 'I am Trump', and by 'I am standing' I say something that won't be true when Trump is standing but DB isn't. The explanation of these facts is that DB is the person whose action *this* idea of action is an idea, and so on. And Anscombe thinks this explanation avoid the various puzzles she has raised, involving immunity to error through misidentification.

But I think it is natural for the referentialist to wonder here why Anscombe doesn't take herself to have just solved all the referentialist's problems as well. If my having ideas of the actions, movements, and postures of DB are enough to make 'I am F' in my mouth an assertion that is true iff DB is F, why can't it also secure the reference of 'I' in my mouth to DB? If each of us has a distinctive idea of the actions, movements, and postures of himself, why can't the conception of oneself that determines the reference of 'I' just be that of the person whose actions, posture, etc. these are the actions, posture, etc. of?

Put more generally, the problems for the referentialist were all supposed to be that there was no unique conception each of us has of ourselves that is guaranteed to apply uniquely to oneself, so as to secure the reference of 'I' to oneself. But Anscombe now grants that we have 'ideas' (and later, 'conceptions') which are uniquely related to oneself so as to secure that the truth conditions of 'I' statements involve oneself. It is natural to wonder why ideas can work to secure the truth conditions of a sentence where modes of presentation failed to secure the reference of a word in that sentence.

Maybe Anscombe imagines some important distinction between modes of presentation and ideas, or between the way that a mode of presentation secures reference to an individual and the way ideas secure truth conditions for 'I' statements. But without any explanation, it's hard to see why the referentialist should grant to Anscombe a way out of the difficulties she raises that isn't equally available to the referentialist.

From what I can tell, here is Anscombe's most direct comment on the matter:

These conceptions are subjectless. That is, they do not involve the connection of what is understood by a predicate with a distinctly conceived subject. The (deeply rooted) grammatical illusion of a subject is what generates all the errors which we have been considering.

I'll admit that I am just at a loss here. Our conceptions of action, posture, etc. are "subjectless", where this just means they do not involve the connection of distinctively conceived subject and predicate. I am not sure if this is true, or why, if it is true, it would prevent the conceptions from fixing the reference of 'I' to the speaker. (Remember, it doesn't prevent them from fixing the truth conditions of the whole sentence 'I am F' so that it concerns the speaker.)

In any case, I think it is fair for the referentialist to complain that Anscombe hasn't sufficiently explained why these conceptions/ideas are ineligible to fix the reference of 'I'. If so, then it makes room for a kind of referentialist view. This view says that there is a special first-personal mode of presentation that gives the meaning and fixes the reference of 'I'. As I've said, I read Anscombe as targeting most of her paper against *this* view, with the circularity problem intended as a problem for the distinct view of simple referentialism.

But now I want to go further, and argue that *if* we grant a special mode of presentation linked to our conception of action, then this opens up a new solution to the circularity problem on behalf of simple referentialism. The first-personal mode of presentation will still figure into the account, but not by giving the meaning of 'I'. Instead, it will only figure in the same way that it does in *all* (or at least many) actions, including ones that have nothing to do with self-reference.

Here is the idea. Suppose Anscombe is wrong about my conception of action being "subjectless". When I intend to walk, that is an intention directed at a particular agent, conceived in a particular way. I intend for *me* to walk. And when I intend to pat some agent X on the back, I intend for *me* to pat X on the back.

Now suppose I intend to pat myself on the back. One account of this intention treats it as like an intention to pat any old person on the back, but with the relevant person being

myself, perhaps conceived of under some distinctive mode of presentation. Consider all the people who have ever patted me on the back. It's the intention to do the thing that all of those people have done—i.e., to pat on the back a particular person, who is me. And then there is the question of how I am conceiving of that person when I intend to pat him on the back.

But maybe there is another way to think of the relevant intention. Consider all the people who have ever patted themselves on the back. Maybe what I intend is best thought of as an intention for me to do the thing that they all did.

To be sure, I don't know the right way to individuate actions here. Maybe the act of doing what all DB-patters do, and the act (performed by me) of doing what all the self-patters do, are the same action. But even so, there seem to be two ways of conceiving of that action. And it seems possible for me to represent the action in my intentions as the one that all the self-patters do. I can just intend to self-pat. I don't need to have some distinct conception of the person I'll be patting, and hope that it turns out to be the same person who is doing the patting.

The idea is that the intention to self-refer using 'I' can be like this. I don't need a distinctive mode of presentation to give the meaning of 'I', in order for it to refer to me. I need a distinctive mode of presentation of myself as the one performing the action, but that is true of many actions having nothing to do with self-reference. Once I have that, my intention can just be for me, the person who acts, to self-refer. It doesn't need to be an intention to refer to an independently identified agent, who let's hope turns out to be the same one doing the referring.

This proposal has something in common with one Doyle discusses from Gareth Evans. Evans' proposal is that the I-speaker's intention to refer to himself can be understood as an intention to bring it about that he satisfies the one-place concept-expression ' ξ refers to ξ '. After some discussion of matters of formulation, Doyle replaces this with the intention to satisfy the open sentence ' x refers to x '. But importantly, both Evans' initial version of the view and Doyle's reformulation have a metalinguistic character. The relevant intention is to bring it about that one satisfies a linguistic or quasi-linguistic construction.

Doyle proceeds to raise an interesting and forceful objection to Evans' view, but I think it might hinge on its metalinguistic character. If so, then something like my own proposal, which is not metalinguistic, despite other similarities to Evans' view, might still stand. Doyle's objection is that satisfying an open sentence is not really an action, so it is not a legitimate object of intention. In contrast, *bringing it about* that one satisfies the open sentence is an action, which one can intend to perform. So Evans could claim that the I-user intends to bring it about that he satisfy the open sentence ' x refers to x '. But now we are back an an intention that involves an indirect reflexive 'he'! For if 'he' is read as an ordinary reflexive, I could have the relevant intention by intending to bring it about that the person spilling the sugar satisfies ' x refers to x ', even if I don't realize that I am spilling the sugar. Although I find it hard to render judgment about this unusual case, it is plausible that I fail to use 'I' with its usual meaning. Thus Evans' account plausibly fails, if the most proximate genuine intention is simply to bring it about that one satisfies the relevant open sentence.

For similar reasons, Doyle thinks we need to distinguish more generally between, e.g., and intention to walk, and an intention to bring it about that I walk, or that I satisfy the open sentence ‘x walks’. Unlike the intention to walk, the latter intentions can be specified only using an indirect reflexive, on pain of the circularity problem. For without the indirect reflexive, they would just amount to an intention to bring it about that a person who is in fact myself walks. This is why, Doyle says, the intention to walk is not covertly just an intention to bring it about that I walk. It is not an intention to bring it about that DB or anyone else, however identified, walks. An intention to walk should not be understood as making tacit reference to myself as the individual whose walking I intend. As we might put it, in intending to walk I am merely the subject of my intention. I don’t make another appearance as its object.

But I’m not so sure this is right. Mundane intentions to go inside, to sit in a chair, to put on a hat, etc. all seem to involve getting this object, this body right here, into relevant relations to other objects. What distinguishes the intention to put on a hat from the intention to see to it that I am wearing a hat is not whether I am conceived of as an object, because I am in both cases, but rather just in the means adopted.

This is important, because if intending to wear a hat involves reference to me as an object—if it is in effect an intention for this person here to put on a hat—then the same can go for an intention to pat myself on the back. It can just be an intention for this person here to pat himself on the back—to do the thing that all the people who pat themselves on the back do. There is thus no need to identify the person whose back will be patted, above and beyond what is involved in even ordinary intentions to walk.

Is my proposal then that an intention to pat myself on the back is the same as an intention to bring it about that I satisfy the open sentence ‘x pats x on the back’? Not really. For one thing, this is because it lacks this metalinguistic character. My intention is to perform an action, rather than to satisfy a metalinguistic condition. For another thing, an intention to ‘bring it about’ that I pat myself on the back is consistent with the use of highly indirect means, such as hiring someone to force me to pat myself on the back. But at the same time, I think the proposal has some of the features Doyle finds objectionable about Evans’ proposal. For it is an intention for “this person here” to pat himself on the back—an intention in which I in some sense figure as an object. But I don’t think this makes the intention any different from many mundane intentions, like an intention to put on a hat, to go inside, to sit in a chair, etc.