

JOHANNES L. BRANDL

## THINKING AND TALKING ABOUT ONESELF

Current debates about self-knowledge often take as their starting point questions about the first-person pronoun.<sup>1</sup> What do I refer to by the pronoun 'I', if I refer to anything by this term, and what propositions do I express by sentences of the form 'I am such-and-such', if these sentences express propositions at all? These questions replace the more traditional questions about self-directed thinking: What is it that I think about when I think about myself, and what is the content of my thoughts when they concern myself?

In this paper I explore the merits of this semantic move. What do we gain by bringing language into the discussion of self-knowledge? In section 1, I suggest that Hume's scepticism about the self gets additional force from a subtle linguistic distinction. In sections 2 and 3 I argue that the two main theories of self-directed thinking, namely the propositional and the non-propositional theory, are unable to meet Hume's challenge in this form. This leaves us with the nominalist alternative according to which thinking is itself an essentially linguistic affair. The question in sections 4 and 5 is whether a nominalist can steer a middle course between a propositional- and a non-propositional theory of thinking. It is from this perspective, I conclude, that the advantages of a linguistic approach to a theory of self-directed thinking must be assessed.

### 1. ME AND MY SELF

In a famous passage in the *Treatise* Hume expresses his doubts about the object of self-reflection:

"... when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other ... I never can catch *myself* at any time without a perception, and never can observe any thing but the perception." (Book I, iv, 6)

Hume's point here can be taken in two ways. On one reading he is arguing for the conception of a self as a mere bundle of perceptions without a substance that exists independently of these perceptions. On a second reading, Hume cuts even deeper and challenges the very idea of a self that can be analysed as a substance, a bundle of perceptions, or whatever.

I want to pick up this more radical interpretation. On this reading Hume's problem is very much a linguistic one. It requires a fine distinction between the term 'myself' as it is used in everyday discourse, and the expression 'my self' as it is used by philosophers. It is good everyday English to say for instance: 'I see myself in the mirror', but not 'I see my self in the mirror'. The latter sentence is philosophers' jargon at best.

In my view Hume rejected the philosophical idiom as an unjustified extension of the way we talk ordinarily. What he demands is that a semantic theory of the first-person pronoun should be as innocent as our everyday discourse. This raises the question where exactly the 'innocence' of everyday discourse ends and where philosophical theorizing about the self begins. I have no recipe for deciding this question, but I would draw the line as follows: The first person pronoun 'I' is used innocently for referring to oneself as long as its referent can be picked out in other ways too, e.g. by a proper name or a definite description. Innocence ends when it is claimed that the first person pronoun is taken to refer to some kind of entity which can be referred to exclusively only by this pronoun.

I should mention here that this way of drawing the line is markedly different from what has been suggested by some interpreters of Wittgenstein.<sup>2</sup> On their view we are facing the following dilemma: either there is something that can be referred to by the pronoun 'I'—and only by the pronoun 'I'—or this singular term should be eliminated. I do not see why anyone should accept this dilemma. That there may be no exclusive referent for the pronoun 'I' provides no reason for avoiding that term completely. Perhaps Wittgenstein feared that the pronoun 'I' makes us susceptible to assuming that such an exclusive referent exists. But this misses the point. What is worrisome is not the pronoun 'I', as it is commonly used for referring to oneself, but the distinctly philosophical use of the term 'self'.

The moral here is that we should not get distracted by discussing whether 'I' is a referential term or not. Common sense allows us to assume that the pronoun 'I' is used as a singular term with a referential function. This still says very little about how this term fits into a general theory of meaning. What we must concentrate on, is the *sense*, not the *reference* of the term 'I'. Here the real problems begin.

## 2. PROPOSITIONAL VERSUS NON-PROPOSITIONAL THEORIES OF SELF-DIRECTED THOUGHTS

There is a basic distinction between two ways of analysing self-directed thoughts, namely the *propositionalist* and the *non-propositionalist* analysis.

Both strategies accept that the first person pronoun 'I' is — at least sometimes — used as a referring term. What these theories disagree about is the following question:

(Q1) Do 'I'-sentences express complete propositions?

Propositionalists (like Castañeda and Evans) give a positive, non-propositionalists (like Chisholm and Lewis) a negative answer to this question. These parties also disagree about another question, however, which is closely related to the first one, namely:

(Q2) Are self-directed attitudes formally like propositional attitudes?

At first, this may seem to be just a different version of the same question. To say that 'I'-sentences do —or do not — express complete propositions may seem to be equivalent to saying that self-directed attitudes must—or must not — be formally treated as non-propositional attitudes. Despite first appearances, however, the two questions are not the same.

As we shall see, it is possible to give a negative answer to question (Q1) while still maintaining a positive answer to question (Q2). This opportunity opens up when we take a closer look at the difference between the propositional and the non-propositional strategies. This difference is best explained by way of an example.

Consider the case of Oedipus who discovers that he himself killed king Laius. The circumstances are such that sentence (1) is true before, sentence (2) only after his discovery:

- (1) Oedipus believes that there is somebody who murdered Laius.
- (2) Oedipus believes that he himself murdered Laius.

How does a propositionalist explain the difference between these two belief-ascriptions? His view is that both sentences are instances of the relational scheme for propositional belief:

(PB)  $B^2(S, P)$

Here 'S' and 'P' are terms designating a subject and a proposition respectively, whereas *B* designates the attitude that *S* has towards *P*, namely believing to be true. Let me emphasize that the term 'P' here must not be confused with the sentential variable in the scheme 'S believes that *p*'. In this latter scheme, what corresponds to the term *P* is the clause 'that *p*' which functions as a nominal designating the second term of the belief-relation.

This formal point turns out to be crucial for solving the Oedipus-problem. The propositionalist must say which proposition Oedipus believes after his discovery. Oedipus himself would express it (if we assume that he speaks English) by saying:

- (3) I am the murderer of Laius.

But Oedipus is dead and we cannot use sentence (3) for expressing what he discovered. Neither will any sentence do in which we replace the pronoun 'I' by some singular term denoting Oedipus:

- (4) *O* is the murderer of Laius.

Whatever we substitute for '*O*' here, sentence (4) will express a proposition that Oedipus may believe without believing that he murdered Laius.

I make a long story short here by concluding that there is no independent sentence by which we could express the proposition that Oedipus discovered to be true. Does that mean that we cannot tell at all what this proposition is? This would be rash. Even if there is no independent sentence by which we can do it, there are sentences in *oratio obliqua* which do the trick. This is the solution of Hector-Neri Castañeda. With the quasi-indicator 'he\*' we can say, but only in indirect discourse, what Oedipus believed, namely that *he\* murdered Laius*.<sup>3</sup>

However, there is also a more straightforward solution. As I said, the propositionalist does not need a sentence to express, but only a term to denote the proposition in question. 'That'-clauses are not the only option, one may also use a description like the following:

- (5) the proposition that Oedipus would express by sentence (3).

A speaker using this description does not actually say what Oedipus discovered, but he nevertheless conveys his belief.

This may help to understand the point Frege makes about Dr. Lauben:

"Wenn nun Dr. Lauben denkt, daß er verwundet worden ist, wird er dabei wahrscheinlich diese ursprüngliche Weise, wie er sich selbst gegeben ist, zugrunde legen. Und den so bestimmten Gedanken kann nur Dr. Lauben selbst fassen."<sup>4</sup>

According to Frege, only Dr. Lauben can grasp the thought that he himself has been wounded. Why should this be so? The explanation that comes to mind is that only Dr. Lauben can *express* this proposition by the sentence 'Ich bin verwundet worden'. That does not mean that Dr. Lauben has some extraordinary linguistic ability, since we all understand the sentence and know

what Dr. Lauben says with it. The point is that we cannot *use* this or any other sentence for expressing the same proposition as Dr. Lauben does.

From this Frege concluded that 'I'-thoughts are not communicable, which is an exaggeration.<sup>5</sup> It is true, however, that they are not communicable in the same way as thoughts which both speaker and hearer can grasp. Here Frege pays his price for a propositional analysis of self-knowledge. His explanation, why certain propositions are not expressible and graspable by everyone, is that these propositions involve some special mode of presentation in which each one is given to himself and only to himself.<sup>6</sup> They involve what Castañeda calls an 'I'-guise of that person.

Surely that will ring the alarm bells of the sceptic. What are these special concepts or guises that are required to explain the peculiarity of first-person propositions? Is there something unique about each of us that each of us knows only by himself? Anyone reluctant to make that assumption should quickly leave the camp of the propositionalist.

### 3. THE ATTRIBUTION-THEORY AND ITS PRESUPPOSITIONS

Roderick Chisholm and David Lewis are the main proponents of a non-propositional theory of intentionality.<sup>7</sup> (John Perry's theory is a different case;<sup>8</sup> I think his views have more affinity with the nominalist position I shall discuss below.) The slogan of the non-propositional theory is: not propositions, but attributes are the content of our intentional states. Attributes can be ascribed in two ways, either to oneself or to an object other than oneself. The difference is formally important. Self-ascribed beliefs (or 'direct attributions', as Chisholm calls them) are two-place relations between a subject *S* and an attribute *A*:

$$(DA) \quad B^2(S, A)$$

By contrast, the scheme for an indirect attribution is a three-place relation between a subject *S*, an object *O* and an attribute *A*:

$$(IA) \quad B^3(S, O, A)$$

The relation between these two schemata raises some complex questions. Chisholm and Lewis take the two-place relation  $B^2$  to be primitive, not reducible to the three-place-relation  $B^3$  by assuming that *S* and *O* are identical. *De se* beliefs are not a special kind of *de re* beliefs, or as Chisholm puts it, direct attributions are not a special kind of indirect attributions. It is just the other way

round because underlying the three-place relation  $B^3$  we only find self-directed beliefs. These are the only beliefs that we *really* have, so to speak. Indirect attributions do not describe different kinds of beliefs, they only describe different relations in which we stand to other things in virtue of having self-directed beliefs.

This view can be challenged in two ways. Either one can argue that *de se* beliefs are in fact reducible to *de re* beliefs, or one can argue that both kinds of belief are primitive and genuine kinds of mental attitude. It is not these possible variants, however, which I want to discuss. I want to look at the consequences of regarding only *de se* beliefs as primitive.

Consider how Chisholm and Lewis would handle the Oedipus-case. They would describe Oedipus' mental state after his discovery as follows:

- (2\*) Oedipus believes himself to have the attribute 'being the murderer of Laius'.

This belief would seem to imply the propositional belief that there is somebody who killed Laius. How do Lewis and Chisholm explain this inference? They do so by reconstructing the propositional belief as a non-propositional attitude as well. Thus the attitude which Oedipus had already before his discovery must be redescribed as follows:

- (1\*) Oedipus believes of himself to have the attribute 'being such that (or 'living in a world in which it is true that) someone murdered Laius'.

This reduction of propositional to non-propositional attitudes has the following peculiarity: It works only if to any proposition  $P$  there corresponds an attribute of the form 'being such that  $p$ '. There is such a corresponding attribute only if for every proposition  $P$  there exists a sentence  $p$  that expresses this proposition for us. In this way, Chisholm and Lewis rule out the existence of special first-person propositions. At the same time, however, they require some special capacities of the subject. She must be able to grasp pretty complex attributes and take herself as exemplifying these attributes. Moreover these capacities are supposed to be primitive:

"We presuppose two things about the abilities or faculties of believers: First, a believer can take himself as his intentional object; that is to say, he can direct his thoughts upon himself. And, secondly, in so doing, grasps or conceives a certain property which he attributes to himself."<sup>9</sup>

The presupposition here is not just that believers have these capacities, but that no further explanation can be given how they are acquired and how they work:

"But to the question: 'What makes his *direct* attribution of a property to himself an attribution of a property to *him*? there can be no answer at all, beyond that of 'He just does – and that is the end of the matter!'"<sup>10</sup>

Will this be the end of the matter for the sceptic too? I doubt it. Surely the sceptic will deny that grasping a property is a primitive capacity. This ability, he will say, rests on our understanding of predicates which express these properties. We can, and must, explain this faculty by explaining how we learn to understand and use the respective predicates.

The same goes for the allegedly primitive capacity of directing ones thoughts upon oneself. This too, the sceptic will say, is an ability we acquire by learning a language, namely by using the first-person pronoun. We can explain what it means to self-ascribe a property by explaining how we learn to understand sentences in the first-person singular. In this way the sceptic demands a nominalistic explanation just where the attribution theory stops.

#### 4. A NOMINALIST THEORY OF SELF-PREDICATION

Quine has marked the route for the nominalist to go. "Instead of speaking of intensions", he says, "we can speak of sentences, naming these by quotation."<sup>11</sup> The propositional scheme gives way to the following nominalistic paraphrase:

(Q) *S* believes-true '...'.  
'...'

This immediately leads back to the Oedipus-problem. The trouble is, as we have seen, that only Oedipus himself can express his discovery by a quotable sentence. Still assuming that he speaks English, Oedipus could say:

(6) I believe-true 'I am the murderer of Laius'.

This is the nominalist paraphrase of 'I believe that I am the murderer of Laius'. But what is the paraphrase of the belief-ascription in the third person:

(7) Oedipus believes that he is the murderer of Laius.

Quine's strategy of quoting the sentence embedded in the belief context gives us the wrong result here. Oedipus does not believe-true the sentence 'He is the murderer of Laius'. At this point the nominalist seems to be forced to fall back on the propositionalist or the non-propositionalist strategies. If he turns to the attribution theory he can say that we ascribe *predicates* instead of *attributes*. This may help to avoid postulating a primitive capacity of grasping properties.

But what about the other capacity, the capacity of directing ones thoughts upon oneself? In this respect the nominalist has gained nothing so far. When he makes use of the primitive notion of 'self-predication' he still raises the suspicion of the sceptic that some dubious notion of a self is involved here.

Thus it seems that the nominalist must pick up the trail of the propositional theory. As we have seen, the propositionalist may use any term for denoting the propositions believed by other subjects, so for instance also the description:

- (5) the proposition Oedipus would express by sentence (3),

where sentence (3) was:

- (3) I am the murderer of Laius.

Consider now the belief-ascription which we get by using description (5):

- (7\*) Oedipus believes in the truth of the proposition which Oedipus would express by sentence (3).

This is a correct instantiation of the propositional scheme (PT). Yet the reference to a proposition does no work here. The nominalist can simply take this analysis and drop the idle part. Thus he gets:

- (7\*\*) Oedipus believes-true sentence (3).

This is unobjectionable on the (counterfactual) assumption that Oedipus is an English speaker. Without this assumption the nominalist must bring in the notion of translation and say: Oedipus believes-true a sentence that is the literal translation of sentence (3). This complicates matters, but it still avoids the appeal to a primitive capacity of self-predication.

Having followed the nominalist this far, it is time to take a critical look at his position. I do not want to rehearse some standard arguments against this view here, because there is a special problem that arises for it in the case of self-directed thoughts.

##### 5. THE PROSPECTS OF A NOMINALIST THEORY OF SELF-DIRECTED THOUGHTS

For a propositionalist the peculiarity of self-directed thinking consists in the fact that the subject can *express* the proposition which she believes and which everybody else can only *denote*. The nominalist, we have seen, makes a



similar move by distinguishing between *saying* and *indicating* what somebody believes. Whereas the propositionalist relies on the expressing/denoting contrast only in explaining self-referential attitudes, however, the nominalist applies this method across the board. When he speaks about somebody's attitudes, he *never* says, but always only indicates what the other one wants, believes, etc. He indicates it by quoting an appropriate sentence. Self-referential attitudes make no exception to this rule.

The nominalist cannot deny, however, that attitudes about oneself pose special problems. For instance, when both Oedipus and his mother believe that Laius is dead they have a belief in common. But when they both take themselves to be guilty they have different beliefs with different truth-values. Where does this fine-grainedness of self-directed beliefs come from?

Attitudes differ either in quality—being a belief, a desire, a hope, etc.—or in content—what is believed, desired, hoped. Given only these two criteria Oedipus and his mother should have the same attitude when they both accept the sentence 'I am guilty'. Yet their attitudes are clearly different.

Can the nominalist solve this problem in the same way as an adherent to the attribution theory does? He would say that Oedipus and his mother have the same belief in one sense (they self-attribute the same property), but in another sense their beliefs are different because they are true under different circumstances. Two persons may share the same belief if we individuate it by quality and content, not however if we include the truth-maker of the belief as an individuating factor.

For a nominalist this would mean to steer a middle course between a propositionalist and a non-propositionalist view. On the one hand he would agree with the propositionalist that there is no *formal* difference between self-referential and other attitudes. On the other hand he would agree with the non-propositionalist that the content of a self-referential attitude goes beyond what is self-attributed or accepted by the subject. What is believed depends in this case on who the believer is.

For the present, I think, we must regard it an open question whether this is a stable position. It is therefore also an open question, I think, whether traditional problems about self-directed thinking can be avoided by the linguistic turn. This leads me to a final, broader observation.

Traditional epistemology since Descartes has tried to exploit the fact that we look at the world through our own eyes, thereby occupying a special position. Anti-Cartesian naturalism has lowered the epistemological temperature. For a naturalist, looking at the world through his own perspective has first of all a practical advantage. By fixing the reference to himself he can refer to other things relative to his own position. This can have practical advantages without requiring any special knowledge about one's own position. One way to make

this point is by saying that self-referential attitudes have nothing special about them: They are neither formally different from other attitudes, nor do they involve special first-person concepts. Viewed in this way they can hardly bear the weight of a Cartesian epistemology.<sup>12</sup>

*University of Salzburg*

#### NOTES

1. For an introduction to the current literature see [Cassam 1994].
2. For a discussion of Wittgenstein's views on the self see [Haller 1989a, 1989b].
3. See [Castañeda 1966, 1968, 1969].
4. [Frege 1918], 39.
5. See [Dummett 1991], 320f.
6. For two (different) attempts at defending this Fregean explanation see [Evans 1981] and [Künne unpubl.].
7. See [Chisholm 1981] and [Lewis 1979].
8. See the papers collected in [Perry 1993].
9. [Chisholm 1981], 28
10. [Chisholm 1981], 32.
11. [Quine 1956], 109.
12. I thank John Bacon for polishing the English of this paper.

#### REFERENCES

- Cassam, Qassim. 1994. (ed.): *Self-Knowledge*. Oxford: Oxford University Press.
- Castañeda, Hector-Neri. 1966. "He: A Study in the Logic of Self-Consciousness." *Ratio* 8: 130-157.
- Castañeda, Hector Neri. 1968. "On the Logic of Attributions of Self-Knowledge to Others." *Journal of Philosophy*, 65: 439-456.
- Castañeda, Hector-Neri. 1969. "On the Phenomeno-Logic of the I." *Akten des XIV. Internationalen Kongresses für Philosophie III*: 260-266.
- Chisholm, Roderick M. 1981. *The First Person*. Brighton: The Harvester Press.
- Dummett, Michael. 1991. "The Relative Priority of Thought and Language." In his *Frege and Other Philosophers*. Oxford: Clarendon Press, 315-324.
- Evans, Gareth. 1981. "Understanding Demonstratives." In his *Collected Papers*. Oxford: Clarendon Press 1985, 291-321.
- Frege, Gottlob. 1918. "Der Gedanke." In his *Logische Untersuchungen*. Göttingen: Vandenhoeck & Ruprecht 1966, 30-53.
- Haller, Rudolf. 1989a. "Bemerkungen zur Egologie Wittgensteins." In: McGuinness, Brian and Rudolf Haller, eds.: *Wittgenstein in Focus*. Amsterdam: Rodopi, 353-373.
- Haller, Rudolf. 1989b. "Unklarheiten über das Ich oder 'Ich', Ludwig Wittgenstein." *Revue Internationale de Philosophie* 43: 249-263.
- Künne, Wolfgang. unpubl. "Frege on 'I'-Thoughts."
- Lewis, David. 1979. "Attitudes *De Dicto* and *De Se*" (with a Postscript). In his *Philosophical Papers*. Vol. I. New York/Oxford: Oxford University Press 1983, 133-159.

- Perry, John. 1993. *The Problem of the Essential Indexical and Other Essays*. New York/Oxford: Oxford University Press.
- Quine, Willard Van Orman. 1956. "Quantifiers and Propositional Attitudes." In Linsky, Leonard, ed. *Reference and Modality*. Oxford: Oxford University Press 1971, 101-111.