

Is Self-Consciousness Consciousness of One's Self?

Jean-Philippe Narboux

1. Introduction

The title of this essay calls into question an equivalence that has been generally taken for granted ever since its terms became available with the advent of modern philosophy. The assumed equivalence holds between self-consciousness, on the one hand, and consciousness of one's self, on the other hand, where one's *self* is by definition that in virtue of which one enjoys self-consciousness. It is built in modern philosophy of mind. I shall call it the *Canonical Equivalence* (hereafter *CE*). Let us have it before our eyes:

Canonical Equivalence:

CE Self-Consciousness is Consciousness of One's Self (SC \leftrightarrow CS)

To the ears of most contemporary philosophers, *CE* sounds perfectly innocuous. What is self-consciousness supposed to be, except consciousness of one's self? *CE* is, in effect, "the first explanation of self-consciousness that may occur to someone, and what the form of the expression [self-consciousness] suggests." (Anscombe 1975, 25; Descombes 2014, 281) It is perhaps only fitting, then, that our question, once rephrased as the question whether consciousness of oneself is consciousness of one's self, should be quite literally inaudible. Indeed, *CE* has the look and the ring of a harmless pun. It is usually assumed before the stage is set, more often than not by default.

According to the greater majority of contemporary philosophers, *CE* can only cease to appear as the truism it is *if* we either deflate the meaning of its left-hand side or (which is more likely) inflate the meaning of its right-hand side, that is to say of the concept of *self*. On this view, qualms about *CE* betray unwitting restrictions on the concept of *self* that make it discordant with the concept of self-consciousness, let alone underwrite it. These qualms are typically traced to the conditional assumption that *if per impossibile* there were such a thing as a "self", it would have to be a Cartesian substantial Ego. They are supposed to evaporate once it is realized that *CE* leaves room for conceiving the *self* as the *aspect* under which a person, in the sense of a living human body, relates to herself, the way she is given to herself.

In this essay, I contend that *CE* is of recent philosophical extraction and that it conceals one philosophically substantial claim, namely, the *Referentialist View of Self-Consciousness* (hereafter *RCS*), as I shall call it:

Referentialist View of Self-Consciousness:

RSC Self-Consciousness is Self-Reference (SC \leftrightarrow SR)

Naturally, if *CE* stands or falls with the *Referentialist View of Self-Consciousness*, we should expect the latter to enjoy the same consideration as *CE* in philosophical circles. And so it does. According to the greater majority of contemporary philosophers, *RSC* can only appear controversial *if* we lay certain

restrictions on the meaning of the term figuring in its right-hand side, thereby inflating it. On this view, qualms about the *Referentialist View of Self-Consciousness* betray unwitting restrictions on the concept of reference that render it unfit to subsume self-consciousness. These qualms are typically traced to the assumption that reference is necessarily receptive in character. They are supposed to dissipate once this assumption is dropped.

Against this view, I argue that *RSC* and *CE* must be rejected as incoherent, as the terms figuring in their right-hand sides are in fact unintelligible. The same holds of the term on their left-hand side, “self-consciousness”, when it is construed in accordance with *CE*. But by contrast with the concept of *self*, which is wholly parasitic on *CE*, the concept of *self-consciousness* is a genuine one, which must be rescued from the demise of *CE*. Doing justice to self-consciousness, far from requiring us to embrace *CE*, requires us to forsake it in even its most innocent-seeming version. The equivalence postulated by *CE* forces the grammar of the first person into the mold of the grammar of the third person and thereby realigns the first person on the third (see Descombes 2014). Accordingly, I maintain against *RSC*:¹

N_R “I”, *qua* manifestation of self-consciousness, is *not* a referring expression.

The anti-referentialist approach to self-consciousness condensed in *N_R* is vindicated in both the works of Sartre and Wittgenstein (especially as read and carried further by Anscombe). In what follows, I draw on the works of these philosophers and try to bring them together in order to motivate my case against *CE*. I take Wittgenstein’s claim that “‘I’ is not the name of a person” (Wittgenstein 1953, §410), Anscombe’s claim that “‘I’ is not a referring expression” (Anscombe 1975, 32), and Sartre’s contention that there is no “‘I’” lying behind consciousness (Sartre 2003b) to articulate a single negative insight into the spuriousness of *CE*. On the exegetical register, my main claim is that their works present far more affinities than it is usually acknowledged, including a shared diagnosis of the source of the equivalence encapsulated in *CE*: namely, that the postulated equivalence rests on a grammatical mix-up, as “it is blown up out of a misconstrue of the reflexive pronoun” (Anscombe 1975, 25).

The striking convergence between the thoughts of these authors has of yet failed to receive the attention that it deserves,² no doubt in part because it undercuts the still pervasive picture of an unbridgeable gulf between phenomenology and analytic philosophy, but also because the negative contention that they share is one that is still poorly understood. In fact, the reflections of Sartre and the later Wittgenstein on self-consciousness seem to have elicited similar responses and suffered similar fates. Both have been charged with indulging in paradox and deemed “extraordinary” (see e.g. Evans 1982). Both have been mistaken for the view that self-consciousness can and must stand aloof from particular human beings. Both have been misunderstood to imply that consciousness is “impersonal” in more or less the sense in which “It is raining” is impersonal.

Bringing together the thoughts of Sartre and Wittgenstein in relation to the topic of self-consciousness reveals the most fundamental divide regarding this topic ascutting the two philosophical traditions known “phenomenology” and “analytic philosophy”.³

2. The Problem of Self-Consciousness

2.1. The Principles of Immunity and Transparency as touchstones for an account of self-consciousness

Self-consciousness is harder to characterize than a default thereof. Thus, William James told the story of a man named “Baldy” who, after falling out of a carriage, made the query: “Who fell out?” and, upon being told that it was Baldy who fell out, cried: “Did Baldy fall out? Poor Baldy!” (See Anscombe 1975, 36) Baldy did not know that *he* had fallen out because, as he was falling out of the carriage, he simply did not realize that *he himself* was doing so. He ended up thinking that *he* had fallen out only insofar as he ended up thinking that *Baldy* had done so. It is worth emphasizing that his companions were able to witness Baldy’s lapse: the difference that self-consciousness makes is *not* a private matter (Anscombe 1975, 36).

That Baldy sought a subject of predication to which to attribute the happening, the falling out of the carriage, attested that his consciousness of the happening came short of self-consciousness. Had his thought of the happening been self-conscious, he would not have made his query. There would have been no need for him to look for a subject, no need for him to identify the person who fell out. What betrayed his lapse was not that he used “Baldy” instead of “I” but that, even though he was conscious and he had the idea that someone had fallen out, he wondered who had done so (Anscombe 1975, 36). Had Baldy phrased his exclamation in the first person (“Did I fall out? Poor me!”), his lapse would have remained conspicuous.

Conversely, although Baldy’s exclamation, “Did Baldy fall out? Poor Baldy!”, has the same content as his antecedent query, “Who fell out?”, and although it further betrays the lapse attending his thought of the happening, it also requires and evinces self-consciousness on his part, simply in virtue of its form, which is that of the explicit performative utterance “I pity Baldy” (see Austin 1975, Chap. 5). The same would hold if he simply had made an assertion like “Baldy fell out”. Indeed, it is far from clear that we could so much as imagine a ‘lapse of self-consciousness’ with respect to the belief that transpires in such an assertion. It would be tantamount to conceiving that Baldy could regard the question whether he had fallen out as settled, yet wonder whether he himself believed that he had fallen out. Which is to say, in turn, that it would have to be open to Baldy to adjudicate the latter question negatively; it would have to be intelligible for him to make the assertion “Baldy fell out, but I don’t believe it”.

Just as a description of one’s motions like “I am falling out” normally leaves no room for a question of the form “Someone is falling out, but is it I?”, so too an expression of belief, say “I believe that Baldy fell out”, absolutely leaves no room for a question of the form “Someone believes that Baldy fell out, but is it I?”. But it is also the case that an expression of belief, say “I believe that Baldy fell out”, absolutely leaves no room for a question of the form “I believe that Baldy fell out, but is it so?” or of the form “Baldy fell out, but do I believe it?” Like seeking to find out whether *I* am undergoing a certain motion that I know to be currently befalling someone, seeking to find out whether *I* endorse the claim that I am making would attest a want of self-consciousness on my part.

If the very need to settle either sort of question shows self-consciousness to be missing, it must be because self-consciousness precludes it from arising. The reason why such an incompatibility obtains is evident: if the need to settle either sort of question could arise when self-consciousness is in place, then it would be conceivable that one hit on the *wrong* answer. But this is absurd.

At this juncture, many contemporary philosophers—including some influenced by Sartre or by Wittgenstein—are likely to invoke the following two principles as the positive grounds of the two negative criteria that were just laid out:

Immunity Principle

IP Self-ascriptions are able to dispense with observation if and only if they are *immune* to error through misidentification relative to the first-person pronoun.

Transparency Principle

TP Self-ascriptions of stances towards the world (beliefs, intentions, or perceptions) are able to dispense with observation if and only if the questions that they purport to answer are *transparent* to corresponding world-directed questions regarding the objects of these stances.

These two principles are meant to spell out two formal features of self-consciousness that respectively account for the two negative criteria. Together, they provide an alternative to the so-called “inner glance model” of self-consciousness which is often ascribed to Descartes (and is almost certainly espoused by Locke). They dislodge the twofold ‘Cartesian’ assumption that the authority of “I” stems from the *incorrigibility* and the *transparency* of consciousness’s epistemic access to itself. Both principles are meant to demystify our ability to make certain first-person claims without resorting to any observation of ourselves (see Boyle, unpublished a).

The two principles do not speak against *CE*. Indeed, they are typically conceived to be not only conditions of adequacy that an account of self-consciousness should meet if it is to avoid perpetuating the “inner glance model” of self-consciousness, but also as conditions of adequacy that an account of *CE* should meet if it is not to lay itself open to the charge of conceiving the self as “a special sort of object”.

The two principles can accommodate *CE* to the extent that they are neutral on these questions: is the reason why the very endeavor to attach one’s conscious motion of falling to the right subject of predication betrays a lapse of self-consciousness, that for a self-conscious being there is no link to be made, or that the link is effortless, guaranteed to succeed? What is wrong with Baldy’s trying to find out whether he is the one who is falling out: that he is looking for a subject when there is none to look for, or that there is no need to look for it? Again, is the reason why the very endeavor to attach one’s first-person assertion “I am falling out” to the right attitude on one’s part, were it genuinely conceivable, would betray a want of self-consciousness, that for a self-conscious being there is no link to be made, or rather that the link is effortless, guaranteed to succeed? What would be wrong with looking for an answer to the question whether I endorsed my assertion “I am falling out”: that I should look for an answer where there is none to look for, or that there is no need to look for it?

But that the two principles can accommodate *CE*, I contend hereafter, does not speak for *CE*. Rather, it speaks against them. As I try to show, they are but the ultimate refuge of *CE* and must likewise be rejected. Unlike the two negative criteria that they pretend to ground positively, they do not characterize self-consciousness, so much as the minimal core of the characterization of self-consciousness implicit in *CE*.

2.2. A dispute over the legacy of Wittgenstein's reflections on "I"

The two principles of *Immunity* and *Transparency* are indirectly rooted in Wittgenstein's work. They were introduced in order to capture two insights into the logic of "I" that his work affords. Whether they actually capture these two insights is however open to question.

Evans is standardly credited with one of the earliest formulations of what has come to be known as the "Transparency Principle" (see Moran 2001, 61). In an influential passage in *The Varieties of Reference*, Evans writes:

Wittgenstein is reported to have said in an Oxford discussion: "If a man says to me, *looking at the sky*, 'I think it is going to rain, therefore I exist', I do not understand him." (...) The crucial point is the one I have italicized: in making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward—upon the world. If someone asks me "Do you think there is going to be third world war?", I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question "Will there be a third world war?" (Evans 1982, 225)

At another place, Evans makes basically the same point about self-ascriptions of perceptual awareness, in terms that are reminiscent of Sartre's original formulation of the principle of intentionality: "What we are aware of, when we know that we see a tree, he says, is *nothing but a tree*" (Evans 1982, 230).⁴

As the quotation above indicates, Evans credits Wittgenstein with the insight that the *Transparency Principle* unfolds. In effect, in the context of reflecting upon Moore's Paradox (i.e. the paradox presented by the assertion "*p* but I don't believe that *p*"), Wittgenstein pointed out that, in at least some occasions, the assertion of "*p*" and the assertion "I believe that *p*" were simply equivalent and traced their equivalence to the asymmetry between the first and third persons of the present indicative evinced by "psychological verbs" (like "believe" or "intend") with respect to the justification of the utterances in which they figure:

That he believes so and so results for us from the observation of his person (*aus der Beobachtung seiner Person*), but he does not make the statement "I believe..." on the ground of self-observation (*auf Grund der Selbstbeobachtung*). And *that is the reason why* "I believe that *p*" may be equivalent to the assertion of "*p*". It is equally the reason why the question "Is it so?" may be equivalent to the proposition "I wish to know whether it is so". (Wittgenstein 1980, §504)

Following Anscombe, I shall call a first-person utterance (by which I simply mean an utterance with "I" as syntactical subject) that is not made on the basis of self-observation (*Selbstbeobachtung*) an "unmediated utterance" or simply an "I-utterance" and the thought that is expressed by it an "unmediated thought" or simply an "I-thought" (Anscombe 1975). As we shall see, Sartre and Anscombe are one in maintaining that where an equivalence of the sort obtains, the unmediated utterance is in no way *about* the utterer. On the view that they share,

the equivalence precludes construing the relevant unmediated self-ascriptions as having the character of *predications* (i.e. as featuring a subject of predication in the logical sense). Evans contests this diagnosis. On his view, what the possibility of such equivalence establishes is only that an I-utterance, in virtue of its being unmediated, *need not* be *expressly about* its utterer. In adjudicating this dispute, one question we shall have to address is whether the *Transparency Principle* is indeed a mere elaboration of Wittgenstein's point, as Evans assumes, or rather a distortion of it, as I think Sartre and Anscombe alike would contend.

The insight underlying the *Immunity Principle* is likewise usually credited to Wittgenstein. In an influential passage in *The Blue Book*, Wittgenstein introduces a distinction between two uses of "I", which he calls "the use as object" and "the use as subject".⁵ In the first use, illustrated by "I have a bump on my forehead", there is room for a certain kind of error for which there is no room in the second, illustrated by "I feel toothache", namely an *error of identification*. Looking in the mirror, I can mistake someone else's forehead for mine. When "I" is used in the second way, i.e. "as subject", I cannot be asked, "How do you know/Are you sure that it is you who are X...?". Nor can I ask myself, "Someone is X..., but is it me?" Thus, not every *de se* thought is an I-thought in this sense. A first-person utterance in which "I" is used "as object" expresses a *de se* thought, but it does not express any I-thought, since it is mediated. In the present indicative, there is an asymmetry between an unmediated first-person utterance (e.g. "I am lifting my arm") and its third-person counterpart ("She is lifting her arm") with respect to their justification: whereas the first-person utterance does not rest on observation, its third-person counterpart does. By contrast, a first-person utterance in which "I" is used as object (e.g. "I have a wart on my wrist") rests on the same kind of grounds as would warrant saying the same thing of someone else.

Anscombe's denial that "I" is a referring expression, unqualified as it may be (she makes no mention of Wittgenstein's distinction between two uses of "I"), is quite evidently a descendant of Wittgenstein's gloss on the use of "I" as subject (McDowell 2009, 187-188; Longuenesse 2017, 49). In effect, Anscombe traces the illusion that "getting hold of the right object is guaranteed" to the observation that "getting hold of the wrong object is excluded" (Anscombe 1975, 32), an observation directly echoing Wittgenstein's remark that, when "I" is used as subject, no provision is made for a certain possibility of error.

The notion of immunity to error through misidentification constitutes another and possibly alternative offspring of Wittgenstein's gloss of the use of "I" as subject. As originally introduced, the notion of immunity to error through misidentification aimed at little more than paraphrase: according to Shoemaker's definition, an utterance in which "I" figures as syntactical subject is "immune to error through misidentification relative to the first-pronoun" if and only if it is "immune to error due to a misrecognition of a person" (Shoemaker 1968, 556).

However, the phrase coined by Shoemaker conveys the implication that the error against whose possibility the use of "I" "as subject" is protected is one that *could* have been logically in place. What is more, the notion is expressly designed to allow driving a wedge between Wittgenstein's characterization of the use of "I" "as subject" and his denial that utterances in which "I" is thus employed are "*about* a particular person" (see Evans 1982, 217-218; Longuenesse 2017, 21). Accordingly, the notion is often presented as an alternative account of the

phenomenon delineated by Wittgenstein (see e.g. Longuenesse 2017, 21) or at any rate as an “amendment to Wittgenstein’s distinction” (Longuenesse 2017, 19; see also Evans 1982, Section 7.2).

Arguably, the notion figures in so many words in Anscombe’s attack on the view that “I” is guaranteed to refer (see Anscombe 1975, 30). There, however, far from elaborating on Wittgenstein’s point, it serves to delineate its ultimate target (see below Section 3.1). To Anscombe herself, it was obvious that the notion of immunity was profoundly Cartesian. The role played by the notion of immunity in Anscombe’s amplification of Wittgenstein has eluded most readers of her essay (see e.g. Evans 1982, Longuenesse 2017). This opens up a possibility parallel to the one we envisioned concerning the *Transparency Principle*: namely, that the *Immunity Principle*, notwithstanding its alleged history, in fact runs counter to Wittgenstein’s insights into the workings of “I”.

Two issues naturally arise in connection with the terminology and the examples deployed by Wittgenstein in the quoted passage. First, it is clear that the question whose nonsensicality concerns Wittgenstein is the question “Someone is F, but is it me?” and not the question “Someone is being spoken of by me by means of ‘I’, but is it me?” In effect, the latter question is logically out of place across the board, whether the first-person pronoun “I” is used “as subject” or “as object” (see Gnassounou 2010, 72). In other words, *no* first-person utterance has logical room for error at the locus of its syntactical subject. To use an anecdote from Anscombe, the bishop caught with his hand on the knee of the lady sat next to him can pretend that he has mistaken her knee for his own, but he cannot pretend that he has mistaken himself for her (Anscombe 1975, 30; Descombes 2004, 152).

However, we would be ill advised to conclude, on the basis of these considerations alone, that “the difference that Wittgenstein spots has to do with immunity to mis-ascription rather than to misidentification”, so that “he was wrong to think of it as making a distinction between uses of ‘I’” (De Gaynesford 2017, 487). Rather, what we need to recognize is that the error of identification whose nonsensicality is registered by Wittgenstein does not concern the use of “I” on its own but its use within the context of an utterance (see Shoemaker 1968, 557; Longuenesse 2017, 49).

But the fact that I can never make sense of the question of *who* is being spoken of by me by means of “I”, no matter how I am using “I”, is significant. For it shows that there is always a logical gap between the thought expressed by an utterance of the form “I am F” and the one expressed by an utterance of the form “The G is F”. There is, in a way, as much of a gap when “I” is used “as object” as there is when it is used “as subject”. In other words, whether a first-person utterance expresses an unmediated or a mediated thought is in that respect immaterial (Anscombe 1975, 23; Evans 1982, 210; McDowell 2009, 189). Although not all first-person utterances are “I-utterances”, i.e. expressions of I-thoughts or unmediated thoughts, self-consciousness manifests itself in each and every first-person utterance.

Second, Wittgenstein’s choice of examples for the first case notwithstanding, the categorial distinction that is captured by the distinction between the use of “I” “as object” and its use “as subject” does not align with the distinction between physical and mental self-ascriptions.

Recall Baldy's lapse of self-consciousness. It showed in his lack of an unmediated conception of the motion (falling out of the carriage) of the person Baldy. In general, one's conceptions of one's own postures, motions, and actions—that is to say, of the postures, motions, and actions of one's own *body*—are usually unmediated, as they are normally not based on the observation of one's body, in contrast to the conceptions that one has of the postures, motions, and actions of other peoples (see Anscombe 1975, 33, 35-36).

A prime example of unmediated bodily self-ascriptions is to be found in the conceptions that one normally has of one's bodily posture, i.e. of the position of one's limbs. The form of knowledge that we normally have (or seem to have) of our bodily posture possesses a special significance on account of two features: first, it does not rest upon observation and therefore is *not receptive* in character; second, it reaches beyond the inner recesses of the mind to a *material configuration*, even though it is knowledge strictly "from inside" (Anscombe 1957 §8, §28). An even more crucial type of unmediated bodily self-ascriptions is constituted by self-ascriptions of intentional actions (Anscombe 1957, §28).

The insight that unmediated self-ascriptions encompass some bodily self-ascriptions is, if anything, equally central to Evans's account of self-consciousness, as elaborated in Chapter 7 of *The Varieties of Reference*. Evans maintains that the phenomenon of immunity to error through misidentification with respect to the first-person is not confined to narrowly mental self-ascriptions but extends to some bodily self-ascriptions and hence to some physical self-ascriptions.

Evans circumscribes three ways in which we can be aware of our own physical states and properties without incurring the peril of error through misidentification.⁶ According to him, the corresponding first-person judgments display immunity in virtue of their being "identification-free": none of them can be decomposed into a pair of judgments of the form < That person is F, I am that person > or more generally of the form < *b* is F, I am *b* > (where "*b*" is schematic for a singular term referring to an object of the relevant kind) (Evans 1982, 221). In contrast with them, first-person judgments in which "I" is used "as object" are "identification-dependent", in the sense that they comprise, and so are mediated by, an "identification-component" of the form < I am *b* > (Evans 1982, 180).

Although Evans's account of the distinction between the use of "I" "as subject" and its use "as object" could not be accepted by Anscombe as it stands,⁷ not only are their explications of Wittgenstein's distinction extensionally equivalent, but also they concur in the view that the distinction is best cashed out in terms of a contrast between mediated and unmediated thoughts. Moreover, as Evans himself notes in Chapter 7 of *The Varieties of Reference*, Sartre is to be credited with this very same insight.

But the examples of I-utterances actually selected by Wittgenstein in the *Blue Book* are all examples of *narrowly* mental first-person utterances, i.e. of *non-physical* unmediated utterances. In effect, they all belong to the subgroup of unmediated utterances that Wittgenstein elsewhere calls "expressive claims" (*Äußerungen*, a term sometimes translated, somewhat misleadingly, as "avowals") and whose defining feature is that they employ what he calls "psychological verbs". The category of these verbs coincides in extension with the category of verbs for *cogitationes* in Descartes's sense (see Descombes 2013; Narboux 2017a). These unmediated utterances can be called "psychological

utterances” provided this designation is not held to imply that they are about the mental states of their utterer (see below Section 3.3). There is a contrast here. Whereas such first-person utterances as “I am lifting my arm” or “I am sitting with my legs crossed” might conceivably be made on the basis of observation, and so fail to express I-thoughts, first-person expressive claims like “I feel toothache” simply cannot be made on the basis of observation and so cannot fail to be unmediated utterances. In other words, “psychological verbs” are such that they necessarily give rise to the asymmetry between the first and third persons in the present indicative.

In effect, that an unmediated utterance is not a claim to know by observation might mean either of two things. It might mean that it is a claim to know otherwise than by observation, or it might mean that it is not a knowledge-claim in the first place. For a first-person utterance like “I am painting the wall in yellow” or “I am standing” to express an unmediated thought is for it to be a claim to know without observation. The point of speaking of “knowledge” here is that it might turn out that I was mistaken in claiming to know (without observation) that I was painting the wall in yellow, or that I was standing. By contrast, the first-person expressive claim “I feel toothache”—which, if truthful, cannot fail to express an unmediated thought—is neither a claim to know by observation nor a claim to know without observation, for the simple reason that it is not a knowledge-claim at all. Accordingly, since my “ability to say” that I feel toothache cannot be contested as false, but only as untruthful, it seems best not to regard it as “knowledge” (see Anscombe 1957, §8).

Whereas Anscombe is silent on the foregoing exegetical issue, Evans declares the passage from the *Blue Book* responsible for having obliterated the possibility for *bodily* self-ascription to be unmediated (or “identification-free in the narrow sense”, to use his terminology). He thinks that the passage has thereby encouraged the false impression that the notion of an unmediated self-ascription is restricted to the mental realm (Evans 1982, 216). In turn, this false impression generates the false notion “that our ‘I’-thoughts leave it open, as a possibility, that we are perhaps *nothing but* a mind.” (Evans 1982, 217)⁸

The standard reception of Wittgenstein’s passage has it that the contrast drawn there “seems to leave it open to use to suppose that at least in the use as object ‘I’ is a referring expression” (McDowell 2009, 187; see also Evans 1982, 217; Longuenesse 2017, 49). On this view, Wittgenstein proposes that “I”, when used as subject, fails to refer, in order to *explain* the immunity to error through misidentification evinced by the utterances in which it is so employed (see e.g. Longuenesse 2017, 21). Anscombe and Evans are then naturally said to “have gone in opposite directions in their accounts of what is unique about the role of ‘I.’” (Longuenesse 2017, 49, 50) *If* the two principles of *Immunity* and *Transparency* can be ascribed to Wittgenstein himself, then Anscombe is indeed going farther than Wittgenstein—and presumably too far. However, I shall contest this account of the dispute (see below Section 3.2).

2.3. Self-consciousness and the so-called ‘fundamental rule of reference’

If self-consciousness manifests itself in each and every use of “I” (and cognate expressions), then “the explanation of the word ‘I’ as ‘the word which each of us uses to speak of himself’ is hardly an explanation” (Anscombe 1975: 22-23;

1976: 55; see also Rödl 2007: 1) given that it fails to provide a non-circular account of self-consciousness. In fact, if “himself”, as it figures in this explanation, is the so-called “direct reflexive pronoun”, then that explanation is surely incorrect. For one may “speak of oneself” (where “oneself” is the direct reflexive), yet not know that the object that one is speaking of is oneself, as Oedipus notoriously did when, lacking the knowledge of the identity of the murderer of Laius with himself, he declared: “The murderer of Laius shall be banished”. Likewise a man named “John Smith”, in speaking of “John Horatio Auberon Smith” (named in a will), may be speaking of himself, yet not know it. In fact, as long as we take the reflexive “himself” as the direct reflexive pronoun, we cannot so much as specify the identity that eludes Smith. For example, it won’t do to say that what Smith fails to realize, is the identity of an object he calls “Smith” with *himself*, since it is still open to construe “himself” here as the direct reflexive (see Anscombe 1975, 23; Anscombe 1976, 55).

For the above explanation of “I” to count as correct, “himself” must be understood, not as the direct reflexive, but as a peculiar reflexive, the so-called “indirect reflexive” (which is by definition the reflexive of indirect speech), as in “Smith did not realize that he himself was named in the will”, or “Oedipus did not realize that he himself was to be banished”. But then the use of the reflexive can only be explained in terms of “I”, so that the attempted definition, while it is not incorrect, nonetheless explains nothing. What we take Oedipus to have failed to realize is: “I am the murderer of Laius”. Likewise what we take Smith to have failed to realize is: “I am Smith”. In other words, the so-called “indirect reflexive” “himself” is but “an *oratio obliqua* proxy for the first-person pronoun of *oratio recta*.” (Geach 1974, 129) The point equally applies to self-consciousness, inasmuch as it manifests itself in the use of “I”. As Sebastian Rödl puts it, “Self-consciousness is properly described as the power to think of oneself only if we lay it down that ‘oneself’, here, is a form of ‘I.’” (Rödl 2007, 1) Thus, the propounded explanation of “I” is either false or circular.

The use of the word “I” evidently conforms to the following rule: “I” designates the one who is using it. The meaning of this rule can be unfolded by means of the following truth-functional equivalence: “If X asserts something with ‘I’ as subject [*i.e.* as syntactical subject], his assertion will be true if and only if what he asserts is true of X” (see Anscombe 1975, 32-33). To speak of oneself by means of the device “I” is, to this extent, to indulge in “self-designation” (see Descombes 2007, 405; Descombes 2009, 150-151).⁹ Plainly, the above rule is not meant as the provision of a synonym for “I”. For “I” obviously does not mean the same as “the one who is using it” (Kaplan 1989a; Kaplan 2005 quoted by Kripke in Kripke 2011, 295; Kripke 2011, 295-296, 302).

It is nonetheless tempting to regard the above rule as providing the core of a semantical account of “I”. It is worth emphasizing, then, that no such rule can account for the use of “I”. It falls short of capturing the use of “I” precisely inasmuch as it fails to capture the self-consciousness that manifests itself in its use (Anscombe 1975, 25; Rödl 2007, 1-4; Kripke 2011, 300-301). It can be shown that it is an error to think that the difference between “I” and an ordinary proper name comes only to this, that everyone makes use of “I” only to speak of himself (Anscombe 1975, 24). In other words, it can be shown to be a mistake to think that the reason why “I have pain” does not stand to “He has pain” as “JPN has pain” stands to “Smith has pain”—or for that matter the reason why “I was

born in France” does not stand to “He was born in France” as “JPN was born in France” stands to “Smith was born in France”—is simply the trivial one that “I” is not an *ordinary* proper name.¹⁰

Another way of putting the same point is to remark that the above truth-functional equivalence could not convey the concept of “I” to someone who lacked it: in order for someone to map the truth-functional equivalence onto *her own case*, she should already possess the concept of “I” (Kripke 2011, 300-301). As Anscombe puts it, “It may be very well to describe what selves are; but if I do not know that I am a self, then I cannot mean a self by “I.” (Anscombe 1975, 29) In recoiling from the view that the use of “I” is made intelligible by the rule that “I” designates the one who is using it, we are likely to fall into the view that at bottom “I” is intelligible only to the one who is using it. Thus, we may succumb to the temptation of replacing the semantical rule for “I” with a “demonstrative theory of indexicals”, as if what determined the referent of “I” were some act of private ostension through which the utterer points to himself in a special inner way (Kaplan 1989, 534-535; Kripke 2011, 299). As in turn the nonsensicality of this view dawns on us, we are likely to fall back into the view that the semantical rule supplies all that is needed.¹¹

Is there a way out of this oscillation? The diagnosis of its source that I shall advocate in the remainder of this essay is summed up in the following passage:

The dispute is self-perpetuating, endless, irresoluble, so long as we adhere to the initial assumption, made so far by all the parties to it: that “I” is a referring expression. So long as that is the assumption you will get the deep division between those whose considerations show that they have not perceived the difficulty – for them “I” is in principle no different from my “A”; and those who do – or would – perceive the difference and are led to rave in consequence. (Anscombe 1975, 32)

This is also, in substance, Sartre’s diagnosis, as we shall see in Part 4.

To summarize, the assumption that self-consciousness can and need be traced to the consciousness of a self, in accordance with *CE*, is rooted in the oscillation between two candidate explanations of self-consciousness, the first of which (in terms of the so-called “indirect reflexive”) adequately characterizes self-consciousness but is not an explanation of it, the second of which (the so-called “fundamental rule of reference”) is an adequate explanation, but not of self-consciousness. The “fundamental rule of reference” articulates the benign sense in which “I” can be said to “refer” (in the standard case, “I” refers to the person out of whose mouth it is coming). But as this rule does not so much as begin to capture self-consciousness (if anything, it presupposes it), it neither supports nor contradicts the *Referentialist View of Self-Consciousness*. Anscombe’s denial that “I” refers is not a denial of this rule. At most, it is a denial of its pretension to be intelligible on its own. As for the symmetrically opposite suspicion that it is a denial of the existence of self-consciousness, it should be clear by now that nothing could lie further from the truth.

2.4. The terms of the *Canonical Equivalence*

In light of the foregoing diagnosis, we are now in a position to flesh out the two limbs of *CE*. “Self-consciousness”, we characterized in the previous section (2.3) as “consciousness that such-and-such holds of oneself”, “oneself” being the so-

called “indirect reflexive”, that is to say, the reflexive of indirect speech (Anscombe 1975, 26). Self-consciousness is what Baldy (in the anecdote told by William James) and the entire company of Alex-users (in the thought-experiment we took over from Anscombe) are conspicuously lacking. It is left out by the so-called “fundamental rule” governing the use of “I”. It is the common feature shared by all *de se* thoughts, whether or not they are I-thoughts, i.e. unmediated thoughts. However, as we saw, our observation that self-consciousness is the character of *that which cannot be reported in the third person except as by means of an indirect reflexive* is nothing more—if also nothing less—than a grammatical reminder. For the irreducibility of the indirect reflexive to the direct reflexive is in fact equivalent to the irreducibility of the first person to the third person. In effect, the employment of indirect reflexive indicators cannot be explained except as by mentioning the internal relation that they bear to the word “I”. Indirect reflexives cannot be characterized except as *oratio obliqua* proxies for the first-person index of *oratio recta*. One cannot so much as ascribe a *de se* thought to someone without making her talk, as it were.¹²

The term “self” is implicitly defined by *CE*, which can be said to give its etymology. As John McDowell once put it:

It is useful to reflect on the etymology of ‘the self’. The self is presumably whatever it is about which a thinker thinks when he thinks about *himself*. (Evans 1982, 259)

As implicitly defined by *CE*, a *self* is either *a certain object* or an *object-under-a-certain-aspect* (i.e. an object as given in a certain way) (Anscombe 1975, 25). By definition, it is what we must *have* or *be* if we are to say and mean “I”, enjoy self-consciousness: it is either that by virtue of *having* which something is endowed with self-consciousness (for example, a Cartesian Ego), or what something *is*, insofar as it is given to itself under the special aspect of self-consciousness (for example, a living human body or person (in a non-Lockean sense) as presented to itself in that way). In the first case, “I” logically functions as a name that each of us employs for her/his own *self*, as distinct from the object bearing the proper name that others call her/him. An “Ego” in Descartes’s sense, a “person” in Locke’s sense, are paradigmatic examples of *selves* as understood this way. In the second case, “I” logically functions as a name that each of us employs for the object that she/he is (i.e. that bears the proper name that others call her/him), insofar as she/he is given to herself/himself under the special aspect that constitutes the sense of “I”. Typically, a *self*, as so understood, is what a living human body or person (in a non-Lockean sense) is, insofar as she enjoys self-consciousness. There is no assumption to the effect that the special way of being given to oneself that defines “I” is qualitative, i.e. that it could be conveyed by means of a definite description. Nor is there any assumption to the effect that this special way of being given to oneself is independent from what is thereby given.

The *modern conception of self-consciousness*, according to which to enjoy self-consciousness is to bear a relation to oneself that is not *contingent* or to relate to oneself *qua* oneself rather than *qua* another, goes far beyond the above characterization of self-consciousness. It implies, and is implied by, *CE*. It is what becomes of the above concept of self-consciousness when it is construed in terms of *CE*. On the present view, the modern conception of self-consciousness, inasmuch as it goes hand in hand with the account of “I” as a referring

expression, is the very one that needs to be relinquished if we are to account for the logical originality of “I”, that is to say, for self-consciousness.

3. Anscombe’s Criticism of the Referentialist View of Self-Consciousness

In this part are presented two arguments by Anscombe against *CE*, both of which are direct elaborations of Wittgenstein’s reflections on the first person. Both take the form of a *reductio* and call into question the very possibility or intelligibility of the equivalence posited by *CE*.

3.1. Anscombe’s first line of thought against the *Canonical Equivalence*

Anscombe’s essay on “The First Person” effects a shift from one philosophical problem to another. The main bulk of the essay (Anscombe 1975, 21-32) is devoted to dismantling a problem that figures at the center of Descartes’s legacy: that of understanding the connection between I-conceptions (i.e. unmediated conceptions) and the object referred to by “I” (see also Kenny 2001, 39-41). This problem is exposed as a pseudo-problem, calling for dissolution rather than for new solutions. Moving away from this problem, the last part of Anscombe’s essay (Anscombe 1975, 33-36) then shifts to another problem, this time genuine: that of understanding the connection between I-conceptions (i.e. unmediated conceptions) and the material object whose states, motions, and actions these conceptions are conceptions of (see also Kenny 2001, 45).

First-person utterances are utterances in which “I” figures as syntactical subject. First-person utterances thus differ from impersonal utterances like “It is raining” in that they do not lack a genuine “subject” in the syntactical sense.¹³ It is of course not in this strictly syntactical sense that first-person are deemed “subjectless” by Anscombe. In maintaining that first-person utterances are “subjectless”, she means to assert that they are devoid of a “subject” in the *logical* sense: such utterances, she holds, lack a subject of predication, that is to say, the first term of a predicative nexus, a target of reference to which to attach a predicate.

As I have indicated, Anscombe’s central claim has been misunderstood in a number of ways. In particular, it has been conflated with the claim that first-person utterances do not evince self-consciousness, or else with the claim that it can make no sense to say that they make a referring use of “I”, when it has not been conflated with both claims at once.

In light of the above, we can readily dispose of these two misunderstandings. First, as we saw in Section 2.3, Anscombe does not deny but on the contrary underscores that the use of “I” manifests self-consciousness. Indeed, in her view, it is precisely this central feature of “I” that precludes it from being assimilable to a device of singular reference. It is precisely inasmuch as it can and must be construed as evincing self-consciousness that “I” need not and cannot be construed as referring to anything—in particular, to anything like a “self”.

Second, as we also saw in Section 2.3, Anscombe does not deny but on the contrary underscores that “I” can be said to “refer” to the utterer, or that to utter “I” is in some sense to “refer to” oneself, if *all* this means is that there exists a truth-value link between a first-person utterance and a third-person utterance

about its utterer. To acknowledge the existence of such a link is to acknowledge that there is, as she puts it, a target of reference “in the offing”. Rather, her point is that this truth-value link can neither capture self-consciousness nor establish that “I” refers to a material substance (see Rödl 2007, 126-127).

Last, but not least, Anscombe’s contention that “I” does not refer is not to be conflated with the claim that “I” refers to nothing (*i.e.* is an empty name), let alone to *a* nothing (*i.e.* a non-entity of sort). In and by itself—that is to say, apart from the fact that it addresses a philosophical confusion—the negative claim “ ‘I’ does not refer” stands on a par with, and is as gratuitous as, the negative claim “ ‘I’ does not climb stairs”. It is simply not privative in character.

The nerve of the main argument of “The First Person” is summed up in this passage:

Getting hold of the wrong object *is* excluded, and that makes us think that getting hold of the right object is guaranteed. But the reason is that there is no getting hold of an object at all. (Anscombe 1975, 32)

Anscombe’s argument, which occupies the central part of her essay (Anscombe 1975, 25-32), divides into two steps, corresponding to the two horns of a single *reductio*. First, she considers and refutes the view that “I” is a proper name (Anscombe 1975, 25-26). Second, she considers and refutes the view that “I”, although it is not a proper name, behaves logically like a proper name and so counts among referring expressions (Anscombe 1975, 27-32). The first step of the argument turns on the objection that if “I” were a name, then we would be faced with the “question what guaranteed that one got hold of the right self” (Anscombe 1975, 25). For the repeated use of “I” would require the re-identification of its reference (Anscombe 1975, 27). The only way of escaping this objection is to give up the claim that “I” is a proper name and to stick solely to the claim that “I” behaves *logically like* a proper name. The second and more important step of the argument hinges on the objection that the escape-view that “I” enjoys “guaranteed reference” “collapses into absurdity when we work it out and try to describe how getting hold of the wrong object may be excluded.” (Anscombe 1975, 31) Thus, the attempt to deflect the demand for criteria of identity turns out to founder into incoherence.

There are essentially two *prima facie* reasons for regarding “I” as a proper name (Anscombe 1975, 26, 29). First, it behaves syntactically like one, since it occupies the same syntactical place (the one we called “syntactical subject”). Second, it can be replaced *salva veritate* by an ordinary proper name X, in virtue of the rule according to which “If X asserts something with ‘I’ as [syntactical] subject, his assertion will be true if and only if what he asserts is true of X” (Anscombe 1975, 32, 29) (see above Section 2.3). As we saw in Section 2.1, the fact that “I” trivially differs from an ordinary proper name in at least this respect, that it is a name everyone uses of oneself (Anscombe 1963, 55; Anscombe 1975, 24) is not enough of a reason *not* to regard “I” as a proper name. For “I” resembles our invented term “Alex”, except insofar as the latter fails to express self-consciousness, and Alex is indisputably a regular proper name (if not one of the ordinary sort) (Anscombe 1975, 26). After all, it was designed as one (see above Section 2.3).

However, the conception of “I” as a regular name of a special kind is faced with an insuperable difficulty. If I am to refer (in the pragmatic sense) to a thing by issuing a token of its name (say “N”), that thing must already carry the name in

question. That is to say, it must have been endowed with that name prior to my use of it; the convention that the name “N” refers (in the semantic sense) to this thing must already be in place. And if this thing is to have been endowed with the name “N”, it must have been individuated by means of criteria of identity. All this holds of “I” if “I” is a proper name. It would seem that some criteria of identity must be at work in the use of “I” if “I” is to refer in the way a proper name does.

This brings us to the true reason for being reluctant to regard “I” as a proper name: if “I” were literally a regular proper name (if not of the ordinary sort), then it would be incumbent upon “I” to identify the referents of its repeated occurrences in the mouth of one and the same speaker as being the same (namely, as being one and the same “self”) (Anscombe 1975, 26-27). In using “I”, I could fail to get hold of the right object, which is absurd.¹⁴

The only way of surmounting this difficulty is to relinquish the claim that “I” is a kind of proper name and to replace it with the weaker claim that it merely displays the *logical behavior* of a proper name, in a word, that it is a device of singular reference. After all, the two considerations initially adduced in favor of regarding “I” as a proper name only pointed to resemblances with respect to logical behavior (Anscombe 1975, 27). Accordingly, the central part of Anscombe’s essay is devoted to blocking the natural if misguided rejoinder that even though “I” does not function as a regular proper name, it nevertheless functions as a device of singular reference in a *bona fide* predication.

We already know that the two considerations initially adduced in favor of the view of “I” as a proper name, although they are correct in themselves, suffice neither to establish that “I” refers in the relatively exacting sense in which a logical subject does, nor to account for the display of self-consciousness that is distinctive of “I” (see above Section 2.3). But the question is whether the logical behavior of a name can be further specified so as to do justice to the specificity of “I”. Put differently, the question is whether self-consciousness can be accounted for in terms of self-reference.

If “I” is a referring expression, then it is “guaranteed to refer” or “secure against reference-failure” in three senses: first, the “I”-user cannot fail to exist so long as she makes use of “I”; second, so long as the “I”-user makes use of it, “I” cannot fail to get hold of an existing and present object; third, so long as the “I”-user makes use of it, “I” cannot fail to get hold of the right object, in other words “I” must be immune to error through misidentification (Anscombe 1975, 28, 29-30). But the only way “I” can refer with all three degrees of immunity is if each use of “I”, presumably each thinking of “I”, guarantees the existence and the identity of its referent by guaranteeing its presence; and the only way each use of “I” (each thinking of “I”), can do so is if it itself defines that of which it guarantees the presence through the very act of guaranteeing its presence. The only way that the use of “I” (the thinking of the I-thought) can identify its referent without appealing to a criterion of identity is if it itself constitutes the criterion of the identity of its referent—if it “has the criterion within himself”, as it were (Anscombe 1976, 51). If “I”, in my mouth, refers to myself, then only *I* can understand how it does. Thus, the view of “I” as a referring expression presupposes the intelligibility of the concept of a private language centered on “I”. This suggests that the fantasy of a private language is at the source of the *Canonical Equivalence, CE*.

In a nutshell, if “I” refers, then “I” either *is* a regular proper name (albeit one of a special kind) or it merely behaves logically like one (that is to say, it is a device of singular reference of another kind). But if “I” is a regular proper name (of a special kind), then “I” is not guaranteed always to refer to the right thing, which is absurd. If on the other hand “I” is not a regular proper name (of a special kind) but rather a referring expression of another kind, then while it might indeed be guaranteed to refer, how it earns this privilege remains altogether unintelligible, which again is absurd. Either way, the consequences are too unacceptable not to cast doubt on the initial assumption that “I” refers.

It is worth stressing, first, that Anscombe’s conditional contention that “If “I” is a referring expression, then Descartes was right about what the referent was” (Anscombe 1975, 31, 32) is not a *premise* of her two-pronged argument, but a by-product of the *reductio* undergone by the idea of “guaranteed reference” (see Descombes 2014); second, that her argument leaves entirely unaffected the grammar of “I”. The charge of paradox laid against Anscombe’s account of “I” rests in good part on a failure to disentangle certain features of the grammar of “I” from *CE*.¹⁵

Our conclusion that “I” is not a referring expression has two important immediate consequences. First, it implies that calling “I” a “first person pronoun” is most infelicitous (as indeed some linguists have long recognized)¹⁶. For the word “I” is not a pronoun and it does not designate the first of three persons. It is not a linguistic expression whose function is to deputize for a noun and so to stand for what a noun stands for. Unlike the so-called “autonomous pronoun” “me” (*Ego, moi*), “I” does not so much as *resemble* a proper name in its employment. “I” refers to a person (in the standard case, the person uttering it, i.e. the person out of whose mouth it is coming) only in the benign sense articulated by the so-called “fundamental rule of reference”. It does not stand on a par with “She/he/it”. It belongs with you, which likewise is not a substitute for a logical subject of predication. The whole notion of a “personal pronoun”, with its attendant tripartition, is spurious. The faulty assimilation of “I” to a pronoun both epitomizes and fuels the philosophical confusion encapsulated in *CE*.

Second, our conclusion implies that utterances of the form “Elsa thinks that she*...” cannot be analyzed on the model of “Elsa thinks that the bus is coming”. They cannot be analyzed as ascriptions of a “propositional attitude”, i.e. of an attitude towards a proposition) (see Castañeda; Geach 1974; Lewis 1979). “Elsa thinks that she*...” means “Elsa self-thinks that...”.

Finally, although the so-called “fundamental rule” governing the use of “I” can neither capture self-consciousness (Anscombe 1975, 25) nor establish that “I” is a device of singular reference (Anscombe 1975, 33), it points towards an all-important question, if only by bringing out the importance of the question “Whose assertion?” (Anscombe 1975, 33):

There is a real question: with what object is my consciousness of action, posture and movement, and are my intentions connected in such 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. (Anscombe 1975: 33)

At the end of Anscombe’s essay, we are left with this new question and with Anscombe’s simple yet undeveloped answer to it.

3.2. A second look at Wittgenstein's *Blue Book*

We can now return to Wittgenstein's *Blue Book* and reconsider the terms of the dispute broached earlier (see above Section 2.2). We just saw that Anscombe introduces the idea of immunity to error through misidentification as the purest expression of the grammatical illusion that gives rise to *CE*. This casts doubt on the claim that this idea is genuinely rooted in Wittgenstein's thought.

Among those who suspect the anti-referentialist approach to self-consciousness of being in the grip of a Cartesian picture of the workings of "I", many take Anscombe's approach to be a radicalization of Wittgenstein's, at least as expounded in the *Blue Book*. On their view, Anscombe pushed Wittgenstein's position in the wrong direction by giving prominence to its anti-referentialist strand. This account of the relation between Anscombe and Wittgenstein's respective stances comprises four tenets. It is held, first, that in the *Blue Book* Wittgenstein rightly admitted that "I" functioned as a referring expression in one of its uses, namely in what he there calls its "use as object" (which involves the recognition of a person) (see *e.g.* Longuenesse 2008, 6); second, that he rightly characterized the contrasting non-referring use of "I", namely, its "use as subject" (which involves no such recognition), as being "immune to error through misidentification" (Evans 1982; Longuenesse 2008); third, that he mistakenly took this characterization of the "use of 'I' as subject" to imply that "I" could not be functioning as a referring expression when it was so used, because he mistakenly took it to preclude "I" from referring to anything else than a Cartesian Ego; finally, fourth, that the anti-referentialist approach articulated by Anscombe in "The First Person" amounts to a generalization of this mistake (Evans 1982). All four claims appear to be wrongheaded.

It must be noted at the outset that the distinction between the use of "I" "as object" and its use "as subject" is introduced for the purpose of exposing the fallacy behind the claim that "I" refers to an immaterial substance. The fallacy takes the following form: the consideration of first-person utterances that do not hinge on the recognition of a particular person by means of her bodily characteristics, coupled with the insistence that "I" serves to recognize something even in such cases, "creates the illusion that we use this word to refer to something bodiless, which, however, has its seat in our body." (Wittgenstein 1957, 69)

But the reason why Wittgenstein singles out psychological utterances for attention is not that he takes them to exhaust the range of unmediated first-person utterances. Rather, the reason is that they coincide with the ones for which Descartes had a predilection. The unmediated thoughts preferred by Descartes, notes Anscombe,

all have this same character, of being far removed in their descriptions from the descriptions of the proceedings, etc., of a person in which they might be verified. (Anscombe 1975: 75)

The Cartesian contention that "I" refers to an immaterial substance specifically results from mishandling the grammar of "the Cartesianly-preferred thoughts" (Anscombe 1975, 35). Not only are Cartesianly-preferred thoughts, like all unmediated thoughts, not based on facts about the speaker, since they do not turn on the recognition of her body by means of her bodily characteristics, but also they are not directly verifiable by attending to facts about the speaker. In

this regard, they are unlike unmediated thoughts that convey non-observational knowledge, like “I am standing”, whose descriptions do coincide with descriptions that are directly verifiable through observation. Wittgenstein is not out to impugn the postulation of a Cartesian Ego because he thinks that if “I” (*per impossibile*) could refer, then it would necessarily refer to a Cartesian Ego. Rather, he proposes that the crucially problematic move is the initial one of holding “I” to perform an act of identification.

In any case, Wittgenstein nowhere suggests that “I” ever plays the function of a target of reference in the sense of a logical subject of predication. There is no suggestion that “I” ever functions as a logical proper name, *i.e.* a device of “singular reference”. As we saw in Section 2.2, *even* first-person utterances in which “I” figure “as object” have no room for error at the locus of their syntactical subject. Only third-person utterances can be wrong at this locus, the reason being that such utterances are predicative connections in the logical sense. Conversely, that first-person utterances do not have room for error at the locus of their syntactical subject is of a nature to suggest that they lack a logical subject. Saying that the use of “I” as object involves on the part of the speaker an act of recognition through which she identifies a particular living human body as being the same as hers—an act of the form <This body is my body>—is not at all the same as saying that it involves on her part an act of ‘self-identification’, *i.e.* an act of identification of herself*. As for the statement “I am this body”, it is simply not an identity statement, as we saw in the previous section (3.1). I conclude that Wittgenstein’s contrast of the two uses of “I” lends no credence to the view that “I” refers in at least *some* of its uses.

The interpretation of the use of “I” “as subject” in terms of its “immunity to error through misidentification” quite evidently goes hand in hand with the interpretation of the use of “I” “as object” in terms of singular reference. If “I” *can* function as a device of identification exposing the speaker to error, then the possibility that it should so function remains conceivable even when it is excluded, since that possibility could have obtained.

3.3. Anscombe’s second line of thought against the *Canonical Equivalence*

A second line of thought impugning *CE* and also stemming from Wittgenstein can be extracted from Anscombe’s work.

We have seen two derivative senses in which an utterance in which “I” figures as syntactical subject might be said to *concern* its utterer, none of which is a sense in which it is *about* its utterer (leaving aside, as before, the sense derived by the so-called “fundamental rule of reference”, which courts equivocation): first, a mediated bodily utterance like “I have a bump on my forehead” can be said to concern me inasmuch as it concerns a particular living human body whom I must recognize by observation to be such that I am it; second, a unmediated bodily utterance like “I am standing” can be said to concern me inasmuch as it concerns this body, whom I am—and here it must be remembered that nothing, in particular no observation, shows me which body is the one (Anscombe 1975, 34). For “I am standing” normally claims to express my unmediated knowledge of my posture—that is to say, the posture of this body, whom I am—and it is directly verifiable or falsifiable in happenings concerning this body, whom I am (Anscombe 1975, 34-35). That my consciousness of posture is “connected with”

this object here, which is my body (Anscombe 1975, 33), however, does not imply that it is *about* me. It does not imply that it is the consciousness of the posture of an object to which “I” refers. “I am this body here” need not be an identity-statement. Indeed, it cannot be one if “I” does not refer.

But if “I” is a referring expression, then *even* expressive claims (*Äußerungen*) concern me, since they are—at least in some way or to some degree—*about* me (where “me” designates whatever object “I” refers to). Conversely, if expressive claims (*Äußerungen*) concern me, then this can only be in virtue of being *about* me, that is to say, in virtue of “I” figuring in them as a device of reference, since they cannot concern me in any derivative sense: thus, for expressive I-utterances to concern me just is for them to turn on acts of self-reference and to be acts of self-predication. But the view that “I” figures in expressive claims as a subject of predication, so that they can be said—at least in some sense or to some extent—to *bear on* the object to which “I” refers, and so presumably on the utterer of “I” (in virtue of the so-called “fundamental rule of reference”), presents two unpalatable consequences.

First, if the I-utterances (i.e. unmediated utterances) that can be grouped under the head of expressive claims (*Äußerungen*) are about an object to which “I” refers, so that they in some way convey how things stand with this object, then they turn on grounds for believing that things do so stand with it, grounds implicating the object to which “I” refers; and no I-utterance pertaining to this class can ever be equivalent in truth-value to an utterance whose content is not tainted by a psychological stance. No utterance whose content is not “subjective” in this sense is ever *detachable* from an I-utterance from this class (see Descombes 2014, 363-364). If “I” refers, then either the two assertions “I believe that there will be a third world-war” and “There will be a third world-war” (and likewise the two assertions “I intend to take my daughter for a visit at the zoo” and “It is good to take my daughter for a visit at the zoo”) are debarred from ever being equivalent in truth-value, as the first in the pair, being an expressive I-utterance, bears at least partly on how things stand psychologically with me, while the second does not; or they *are* equivalent, but only because the second, despite the appearances, *also* bears on myself, if only tacitly. Either way, no ‘objective’ statement can ever be *detached* from an expressive claim (see also Descombes 2002).¹⁷

In any event, the asymmetry that expressive claims present with their third-person counterparts precisely consists in the fact that the former do not turn on certain grounds *why* things stand in a certain way with the speaker. Either it makes no sense for me to adduce grounds for my expressive claim (what could be my grounds for saying “I have a headache?”), or it *does* make sense, but then the grounds in question do not have to do with me, but only with the world (my grounds for asserting “I believe his story” are grounds tending to show that the thing believed is true; my grounds for asserting “I want an apple” are grounds for wanting an apple).¹⁸

Second, if the I-utterances (i.e. unmediated utterances) that can be grouped under the head of expressive claims (*Äußerungen*) are about an object to which “I” refers, then a certain gap cannot fail to open between the utterer of “I” and the object to which “I” refers within his utterance (Descombes 2014, 298). A distance will inevitably arise between the “speaking subject” and the “subject” of whom it allegedly speaks, that is to say, between the agent of the speech-act performed in

and by uttering the I-utterance and the referent of the subject of predication (“I”) allegedly figuring in the uttered sentence.

The second line of thought that was examined in this section invokes against *CE* the second of the two negative criteria of self-consciousness (see above Section 2.1). But this hardly shows that it invokes the *Transparency Principle* as formulated above. For if Anscombe is right, what the principle presents as a transition between two things is no transition at all, as there is only one thing.

The second line of thought impugns the logical *possibility* of the “self”, like the first one. However, it is clear that it can also be seen as impugning its logical *necessity*. Emphasizing the logical impossibility of the “self” and emphasizing its logical superfluity possess symmetric advantages and disadvantages. In emphasizing its logical impossibility, we run the risk of making the rejection of *CE* appear privative and therefore deflationary in character: it may look as if there were some intelligible possibility that we are denying. Hence the impression that Anscombe’s conclusions are “extraordinary” (Evans 1982) or “dramatic” (Longuenesse 2017). Showing the “self” to be unnecessary can free us from this misimpression. Symmetrically, in emphasizing the logical superfluity of the “self”, let alone its “absence”, we run the risk of making the rejection of *CE* appear relative to some prejudicial views regarding the function that the “self” is supposed to fulfill and the manner in which it is supposed to be presented in consciousness. On the second score, it will be objected that it should come as no surprise that the “self” should neither call for prospection nor allow for inspection in the way other objects do. The philosophical idea of the “subject” is After all the idea of “something that is indeed not presented, because it is what presentations are made to”, “an area of darkness out of which light shone on everything else” (Anscombe 1976, 56). The “subject” is, in a word, “invisible” (Anscombe 1975, 32), it necessarily eludes consciousness, as consciousness turns back to get hold of it, because it lies “behind” consciousness (Sartre 2003b, 122). Thus we are led back to the necessity of bringing out the inconsistency of the concept of the “self”.

This oscillation suggests that we must show the impossibility and the superfluity of the “self” to be functions of each other. What must self-consciousness be, so that it can have *nothing* to do with self-reference?

4. Sartre’s Criticism of the Referentialist View of Self-Consciousness

Throughout his philosophical evolution, Sartre rejected *CE* as unintelligible.¹⁹ Just as Anscombe denies that self-consciousness can be explained as a special, reflexive sort of reference, in the same way Sartre denies that self-consciousness can be explained as a special, reflexive sort of intentionality.²⁰ Notwithstanding some irreducible differences between their philosophical outlooks, both authors share this negative diagnosis: the logical structure of self-consciousness stands at odds with the structure exhibited by a predicative nexus. As I shall now try to show, Sartre’s negative findings are not anymore paradoxical than Anscombe’s.²¹

4.1. Sartre on the non-thetic character of self-consciousness

The only place where Sartre argues explicitly and directly against *CE* is in his long early article from 1936, *The Transcendence of the Ego*, whose eponymous thesis he was to regard as all but definitive. There, Sartre argues against *CE* along all three modalities. *Being and Nothingness* neither repeats nor expands on these arguments, but simply takes them for granted. Certainly, it does not mitigate them.²² Contrary to a common assumption, we should take Sartre at his word on this score (see Sartre 1980, 142).

But although *Being and Nothingness* does not argue directly against the notion of an immanent Ego and it rests content with summing up the central thrust of the earlier essay (Sartre 1980, 142), it argues for a thesis that is closely related to, and in fact entails, the thesis of the incoherence of the notion of an immanent Ego: namely, the thesis that the structure of self-consciousness is radically at odds with the “intentional structure” that characterizes all consciousness, the structure by virtue of which all consciousness is “of” something, or “directed at” something. In doing so, the book as a whole secures the unity between the three arguments of the essay and it puts them on firmer foundations.

Since this fundamental point is forcefully made in the third Section of the Introduction to *Being and Nothingness*, at the risk of anachronism, I shall first consider this text before turning to Sartre’s arguments in the early essay. Sartre brings out the following dilemma: on the one hand, my consciousness of something cannot so much as exist without being somehow conscious “of” itself; on the other hand, it cannot be primarily conscious “of” itself in the sense in which it is conscious “of” the thing “at which” it is “directed”.

Sartre argues by *reductio* against the claim that my consciousness is primarily conscious “of” itself by directing itself at itself (Sartre 1980, 18-19). Like Anscombe’s first argument against *RSC*, this argument by *reductio* turns on the rejection of the intelligibility of a subjective reflective diathesis, i.e. an act of reflection by which a subject relates to herself, takes herself as object, but *qua* subject.²³ For my consciousness of a thing to be conscious “of” itself in the way it is conscious of that thing, namely by directing itself at itself, would be for it to *know* itself to be conscious of that thing, or at any rate it to *reflect* itself as conscious of that thing. As the knowledge relation imports a distinction of subject (knower) and object (known), this means that self-consciousness would at once divide into subject and object and overcome its own division. But then it would be a question how consciousness, as knowledge of itself, is guaranteed to be about itself (see Gardner 2009, 46). Sartre concludes that in its primary mode the consciousness that all consciousness has “of” itself cannot be a knowledge or quasi-knowledge (reflection) of itself. The structure of self-consciousness (*conscience de soi*) is not that of self-knowledge, i.e. knowledge of oneself* (*connaissance de soi*) (see also Sartre 2003d, 148-149).

Our predicament is that all consciousness of something must perforce be conscious “of” itself, yet it cannot without inconsistency be directed “at” itself. But how on earth can my consciousness *of* something be self-conscious otherwise than by being conscious *of* itself? How can it be self-conscious otherwise than by reflecting upon itself? An answer is broached on the basis of the well-known example of the act of counting the twelve cigarettes left in a pack of cigarettes (Sartre 2003a, 9).

So long as I was busy counting the cigarettes and I had not yet completed the action of counting them, and thus not yet found it to be the case that “there are

twelve”, my consciousness was directed at the world—at *these* cigarettes-to-be-counted—just as the conclusion that I reached is about the world. Yet, had I been asked, “What are you doing?”, I would have *spontaneously* made the reply, “I am counting”, without so much as interrupting my action of counting the cigarettes and so without ceasing to be absorbed in my counting of them. This spontaneous reply attests not just to my ability to reflect upon what I am doing as I utter these words, but also, and more importantly, to my having been able all along to utter them in virtue of having been *all along unreflectively aware “of”* counting those cigarettes. Although or precisely insofar as it was directed at the cigarettes to be counted and towards the completed action, and thus did not *know* itself to be counting, my consciousness “of” counting was all the while conscious “of” itself*. While I am counting the cigarettes, my consciousness “of” counting does not “posit” itself or takes itself as its own object in the way it “posits” the object of which it is conscious. Rather, it is “non-positionally” or “non-thetically” conscious “of” itself. Thus, the view that *reflection* (i.e. thetic consciousness of consciousness) is the primary or default mode of consciousness “of” consciousness is not simply an impossible one, but also an unnecessary and *de facto* unwarranted one.

If now we probe further the relation between the two principles of intentionality (that all consciousness is thetically conscious of something) and self-consciousness (that all consciousness is non-thetically conscious of itself*), we shall find that it is one of duality. On the one hand, all thetic consciousness of something is non-thetically conscious of itself*. On the other hand, all non-thetic consciousness of oneself* is thetically conscious of something.

The task becomes that of understanding how intentionality and self-consciousness can be at once absolutely distinct and absolutely inseparable or, rather, be the one in virtue of being the other. Sartre underlines that it is really one and the same consciousness that determines itself “in a single stroke” (*d’un seul coup*) as non-positional consciousness of itself and as positional consciousness of an object, so that one’s spontaneous consciousness of one’s consciousness is strictly “one with” (*ne fait qu’un avec*) the latter (Sartre 1980, 20).

The thetic “of-structure” of intentionality and the non-thetic “of-structure” of self-consciousness are not to be conceived as two species of a genre. The use of one and the same sign (“of”) to symbolize these two structures is potentially misleading insofar as it blurs what is an absolute categorial distinction. A more perspicuous notation will resort to two distinct signs. To forestall any risk of equivocation, Sartre famously inserts parentheses around the “of” that belongs with non-thetic consciousness and reserves the unadorned “of” (sometimes italicized) to thetic consciousness, that is to say, to intentionality. Thus, he systematically writes “consciousness (of) itself/oneself” (*conscience (de) soi*) to mark the non-thetic character of such consciousness and to contrast it with “consciousness of itself/oneself” (*conscience de soi*) which, being positional, presupposes reflection.

In the terms of our previous discussion, we can say that the parentheses around the “of” indicate that the reflexive appended to it—the word “self” and its composites—must be understood to be the *indirect reflexive* and *therefore not to be a device of singular reference*. Using as before Castañeda’s star-symbol “*” as a specific marker of the indirect reflexive, we might say that consciousness (of)

itself/oneself is necessarily consciousness of itself*/oneself*. The point of Sartre's typographical innovation is to make perspicuous a logical distinction that the syntax of ordinary language elides. In English, the same notational purpose can be attained, with less artifice, by the ante-position of the reflexive indicator, as in the phrase "self-consciousness", and more generally by the ante-position of first-person indicators, as in the phrases "I-thought" or "I-sentence". This is of course the convention to which this essay has resorted all along.²⁴

At the beginning of *The Transcendence of the Ego*, Sartre asks whether there is place for a referent of "I" in a first-degree or unreflected consciousness of something, given the way in which such a consciousness is conscious of itself, namely non-thetically. His answer is without appeal: "of course not." (Sartre 2004, 8) To secure this negative conclusion, Sartre successively challenges the necessity, the possibility, and the actuality of an immanent referent of "I" lying behind every consciousness. Which is to say that he successively challenges the necessity, the possibility, and the actuality of *CE*. He concludes this three-pronged argument as follows:

There is no I [*Je*] at the unreflected level. While I am running to catch the tramway, while I am looking at the hour, while I am absorbed into the contemplation of a portrait, there is no I. There is consciousness of *the-tramway-having-to-be-caught*, and so on, and non positional consciousness of consciousness. In fact, I am then immersed in the world of objects, *they* constitute the unity of my acts of consciousness, present themselves with values, with attractive and repulsive qualities, but *me* [*moi*], I have vanished, I have annihilated myself. There is no room for *me* [*moi*] at this level, and this does not stem from an accident, from a momentary lapse of attention, but from the very structure of consciousness. (Sartre 2004)

This passage is echoed by a passage from the lecture "Conscience de soi et connaissance de soi" (2003d) that strongly suggests that for Sartre the thesis that one's consciousness of an object is *non-thetically* conscious of itself* and the thesis that it has *no room for an I* (at least on its own) go hand in hand with each other, so that they cannot be understood apart from each other:

There is consciousness of *oneself* [*conscience de soi*] with an underlined "of" in the case where we have a reflective consciousness of ourselves. If on the contrary we consider that I ignore at this moment that I exist, that I am so absorbed that, when I will be removed from my reading I will ask where I am, but that my reading may nonetheless imply the consciousness of my reading, then the consciousness of my reading cannot posit itself, as the consciousness of the book standing before me posits it. (Sartre 2003d, 149)

We are reminded at once of Wittgenstein's remark (quoted above in Part 2 Section 2) that "If a man says to me, *looking at the sky*, 'I think it is going to rain, therefore I exist', I do not understand him." (Evans 1982, 225) The unmediated character of our knowledge of our own intentional actions, i.e. the fact that it does not rest on "self-observation" (*Selbstbeobachtung*) (to use Wittgenstein's phrase, quoted above), is reflected by the possibility of eliding the first-person indicator in stating them—a possibility which, far from impairing the ability of our statements to display self-consciousness, hangs on it (for what shows forth need not be said):

Say I go over to the window and open it. Someone who hears me moving calls out: What are you doing making that noise? I reply "Opening the window". (...) But

I don't say the words like this: "Let me see, what is this body bringing about? Ah yes! The opening of the window". Or even like this "Let me see, what are my movements bringing about? The opening of the window (Anscombe 1957, §28, 51; see also §22, 35)

Sartre's standard examples evince a striking congruity with Anscombe's. When it comes to the topic of intention, her examples typically feature pairs of statements of intentional actions cast in the first and second persons of the present progressive (which is a morphological marker of the imperfective aspect), say, "Why are you pumping?"—"I'm replenishing the water-supply." (Anscombe 1957, §23, 38-39; see Thompson 2008, Chapters 5 and 8; Thompson 2011) Sartre relishes using the equivalent formula in French, in which the imperfective aspect is marked periphrastically.²⁵ This bias is methodologically motivated. It aims at counterbalancing the prevailing bias in favor of "Cartesianly preferred I-thoughts", which has imposed "a diet of examples." We are apt to forget that one way to express an intention is by giving a wider description of what one is doing (Anscombe 1957, §22, 34-35). These examples are meant to free us from a philosophical picture according to which to know someone's intentions is to be apprized of the contents of her mind (Anscombe 1957, §4, 9); a picture that would have us "find in every genuine straightforward rationalization a movement from inner to outer, from mind to world, from 'desire' to 'action'." (Thompson 2008, 90)²⁶

However, the distinctive focus of Sartre's examples is on *absorptive activities*, a paradigmatic example of which is the activity of reading a story (Sartre 2003b, 100-101; Sartre 2003d, 149-150). The hallmark of an absorptive activity is that, while I am engaged in it, I "forget myself", as we like to say, and my surroundings alike (see Fried). As Dan Zahavi puts it:

As long as we are absorbed in the experience [of reading a story], *living* it through, no Ego appears. It is only when we adopt a distancing and objectifying attitude toward the experience in question, that is, when we reflect upon it, that an Ego appears. (Zahavi 2005, 100-101; see also 34-35)

However, Sartre does not regard such evocations as compelling by themselves. They are not meant to supplant argumentation. The evocation of the absorptive action of running to catch the tramway, in the above-quoted passage, is meant to illustrate Sartre's three-pronged argument, not to replace it. He is quite aware of the perils incurred by any attempt at showing the absence of something (see De Coorebyter 2000, 253). Moreover, absorptive activities are only meant to bring into relief the *ontological absorption* that characterizes all consciousness of something.

Fortunately, the convergences between the two authors extend to their arguments, in particular, to their arguments against the intelligibility of *CE*.

The first argument adduced by Sartre in *The Transcendence of the Ego* against the possibility of *CE* (i.e. his first argument that a consciousness, *qua* self-conscious, cannot have room for a self, that the "I" is, as he puts it, destructive of (*nuisible*) self-consciousness) (Sartre 2003b, 98-99; see also Sartre 2003d, 150-151) is consonant with the second argument that I extracted from Wittgenstein and Anscombe in the previous section.

A "subject" lying behind every consciousness, conscious of itself *qua* "subject" would be, to use Husserl's apt formula, a "transcendence in immanence". In this sense, an "I", no matter how "formal" or "pure" it might be, would necessarily be

an *inhabitant* of consciousness (Sartre 2003d, 151). Even if it is infinitely contracted, like a geometrical point, an “I” constitutes “a center of opacity” within consciousness (Sartre 2003b, 98). But consciousness is “all-lightness, all-transparency” (*toute légèreté, toute translucidité*) (Sartre 2003b, 98), it is literally *nothing* but the act of positing its object in front of itself:

It gets conscious of itself *only insofar as it is* [en tant qu'elle est] *conscious of a transcendent object*. Everything is clear and transparent [*lucide*] in consciousness: the object is in front of it, with its characteristic opacity, but it, on the other hand, is purely and simply consciousness of being consciousness of this object, this is the law of its existence. (Sartre 2003b, 98)

A pure “I” would ruin the transparency that is the very mode of existence of consciousness *qua* intentional, as it would necessarily stand in its way. It would, as we shall see in Section 4.3, destroy its constitutive polarization (namely, $\cdot \rightarrow +$). Furthermore, Sartre’s argument against the necessity of *CE* (i.e. his argument that a consciousness need not have room for a self in order to be self-conscious, that the “I” is, as he puts it, “superfluous”) (Sartre 2003b, 96-97) also finds a counterpart in Anscombe’s work. In effect, the first of the three arguments adduced by *The Transcendence of the Ego* is directed against the claim that only an “I” can confer to a consciousness of something the unity and individuality without which that consciousness would disintegrate into thin air (on this argument, see Gardner 2009, 22; Longuenesse 2017, 46).

The parallelism can be taken further. According to Sartre, when I am running to catch the tramway (or reading this novel), my consciousness wholly transcends itself *in (dans)* the tramway-to-be-caught (the novel-to-be-read) and *towards (vers)* my catching of the tramway (my finishing the novel). In other words, my consciousness directs itself receptively *at (sur)* an object of perception (disclosed in light of an end of mine) to which I can egocentrically refer by means of a demonstrative and practically *towards (vers)* the finite end that I am currently pursuing (the possibility of achieving which is disclosed through perception). While I am thus engaged and absorbed *in* the intentional action of running, I am *non-thetically* conscious (of) it. This means that the consciousness that my intentional action of running has (of) itself is “unmediated” in Anscombe’s sense. Indeed, it does not rest on the “quasi-observation” which, for Sartre, defines “reflection” (see Longuenesse 2017, 50, 62).

But if my awareness (of) my intentional action of running (or of reading this novel) cannot be *thetic*, at least so long as I am engaged and absorbed in the action, then neither can the mode of bodily awareness that it involves. Like the former, the latter must be *non-thetic*. Thus, while I am running, I am *non-thetically* aware (of) my legs; while I am writing with a pen, I am *non-thetically* aware (of) my hand.

Sartre calls this *non-thetic* mode of bodily awareness the “body-for-itself” (*corps pour-soi*). To recognize that my *non-thetic* consciousness (of) being *thetically* conscious *of* something, or self-consciousness, generally involves a *non-thetic* consciousness (of) my* body, is to recognize that my default relation to my body is one of being *non-thetically* conscious (of) myself, that is to say, a mode of self-consciousness. This mode of self-consciousness, i.e. the *body-for-itself*, need not be the only one. Nor is the *body-for-itself*, of course, the only relation that I can bear to my own body. I can for example see or touch, and in that way be *thetically* conscious *of* my own body (that is to say, *of* this thing here, that I am).

But if it is true that for my “unmediated agent-or-patient conceptions of actions, happenings and states” of this body to be “unmediated” is for them to be conscious (of) themselves, then it is for them to partake of the non-thetic mode of bodily awareness that Sartre calls the “body for-itself”.

4.2. How to reflect the unreflective. A methodological predicament and its solution

Against the foregoing parallelism between Sartre and Anscombe’s respective accounts of self-consciousness, it is natural to object at once that it overlooks a glaring divergence: Sartre consistently admits, whereas Anscombe emphatically denies, that “I” is a referring expression. Indeed, Sartre expressly countenances that each of us is “intimate with” her own “Ego”. His claim is not that “I” does not refer to an object, but that the object to which it refers is not immanent but transcendent, like any object.

This objection, however, is shortsighted. First, it misses the heart of the matter. The crucial point is that for Sartre self-consciousness is not underwritten by the consciousness *of* one’s Ego. Saying that the Ego is a by-product of reflective consciousness is not at all the same as saying that it is the correlate of self-consciousness. Sartre concurs with Anscombe’s contention that insofar as it manifests self-consciousness, “I” is not, need not, and cannot be in the business of referring. On the one hand, my consciousness of myself*, as non-thetic consciousness (of) myself, is not “directed at” anything, it has no “correlate”. Self-consciousness is not a relation. On the other hand, my Ego, far from lying at the center of all my thoughts, is only the by-product of some of them (those that are instances of impure reflection) and it belongs wholly to the world without me, where it stands on par with all the objects at which my consciousness may be directed. Thus, not only the concept of self that Sartre countenances under the name of “Ego” is not the same as the one that figures in *CE* and that Anscombe impugns, but also it is part of his own criticism of *CE*.

Second, as early as in *The Transcendence of the Ego*, Sartre readily maintains that “I” can figure in the formulation of an unreflected thought, without undermining the non-thetic character of its consciousness of itself* (as shown in its formulation) and thereby carrying it onto the reflective level (Sartre 2004, 40; Longuenesse 2017, 47, 50). This is possible because “I” can figure in the formulation of a thought without actually referring to the self (“Ego”) of the speaker. “I” then functions as an “empty concept”, it has a sense but no referent:

It is however certain that the *I* appears on the unreflected level. If I am asked, “What are you doing?” and I reply, all busy, “I am trying to hang up this picture”, or “I’m repairing the rear tire”, these sentences do not transport us onto the level of reflection, I pronounce them without ceasing to work, without ceasing to have in view exclusively the actions insofar as they are done or to be done—not insofar as I am doing them. (Sartre, 2004, 40)

Conversely, their being expressible by I-utterances does not debar “unmediated agent-or-patient conceptions of actions, happenings and states” in Anscombe’s sense (Anscombe 1975, 36) from being “self-conscious” in Sartre’s sense, i.e. non-thetically conscious of themselves*. I can spontaneously declare, “I’m falling!” (what Baldy might have done, had he not suffered a momentary lapse of self-consciousness), “I’m being carried across the room”, or “I’m going upstairs to

fetch my camera” without leaving the unreflected level (see Longuenesse 2017, 62).

In particular, the proposal that intentional actions are ones to which the question “Why are you *X*-ing?” has application (Anscombe 1957, §§5-6; Rödl 2007, 44), which question evidently calls for an essentially *first-personal* form of explanation—in the paradigmatic case, “I am *X*-ing because I am *Y*-ing” (see Anscombe 1957, §23, 38; Rödl 2007, 55-63; Thompson 2011, 206)—does not imply that the unmediated, non-observational knowledge that their agents have of these actions is not located on the pre-reflective level.²⁷

The point is not that the series of questions “Why?” is anyway only an artificial device (like Aristotle’s series of practical syllogisms) to bring out the order of intentional action, since such a series “cannot occur very often” (Anscombe 1957, §42, 80), but that giving an answer to the question “Why are you *X*-ing?” (e.g. “I’m taking the 7:51 bus to Mehringdamm” in reply to “Why are you running?”, or “I’m fetching my camera” in reply to “Why are you going upstairs?”) need not disrupt one’s absorption in one’s ongoing intentional action.²⁸

Third, the possibility that the word “I” (or at any rate a first-person inflexion, like the “-o” ending in Latin) may occur otherwise than as the outcome of reflection appears to be analytically contained in the concept of a “pre-reflective *Cogito*”, if the latter is not to be reckoned ill-named or ineffable. If there is room for a “pre-reflective *Cogito*”, then there is room for a subjectless *Cogito*. The above objection nevertheless touches on a genuine difficulty. In a nutshell, the difficulty is this: how can the non-thetic character of self-consciousness be the subject of philosophical elucidation if philosophy is an essentially reflective, hence thetic, activity? (See De Coorebyter 2000, 351) To understand the nature of this difficulty, it is necessary to take a closer look at Sartre’s assessment of the *Cogito* in *The Transcendence of the Ego* (2003b, 99-104). We shall thereby get in a position to examine Sartre’s third and last argument against *CE* in that essay, namely his argument that no “I” can be found in self-consciousness.

As Sartre notes, Kant’s thesis that “The ‘I think’ must *be able* to accompany all my representations”, given its stress on the embedded modality (that of possibility), does not appear to be meant to settle the *de facto* question of whether the “I think”, in point of fact, does accompany any representation of ours.²⁹ In contrast, the performance encapsulated in Descartes’s *Cogito* provides an effective counterpart to Kant’s thesis, an effective route to the conclusion that the “I think” can be prefixed to any representation of mine. But the crucial question whether in doing so the Cartesian performance settles the *de facto* question that is Sartre’s concern proves a tricky one. Sartre argues that, contrary to what one might have hoped, the reflective *Cogito* does not, because it cannot, settle that question.

Suppose that, only an hour ago, I was absorbed in reading this novel, and that I now recall having been reading. I can recall *what* I was then reading as well as the *circumstances* in which I did. But I can also recall *my* having been reading this novel, that is to say, I can recall that it was *I* who was then reading it. In other words, I can perform any recollection “on the personal mode” (Sartre 2003b, 99). In that sense, the “I think” can indeed be prefixed to all my representations. But does my recollection “on the personal mode” establish that, while I was reading, I was already aware, not just of doing so, but also of *my* doing so, i.e. of *myself* doing so? The question is in fact fundamentally equivocal as it stands.

It is worth dwelling on the reflective structure of the Cartesian *Cogito* (Sartre 2003b, 99-100). "I think" (= *cogito*), insofar as it is directed at or posits another consciousness, namely my consciousness (of) thinking, gives expression to a reflecting consciousness. This reflecting consciousness is non-thetically conscious (of) itself, just as the unreflected consciousness that it posits was non-thetically conscious (of) itself before it was subjected to reflection and thereby made into a reflected consciousness. So that I became conscious of my reflecting consciousness, a third-order consciousness directed at my reflecting consciousness would be required. This means that "I think", although it gives expression to my *reflecting* consciousness of my consciousness (of) thinking and it speaks about my consciousness (of) thinking, is not the expression on my part of a consciousness's reflective consciousness of itself* (Sartre 2003b, 100). "I" and "think" never stand on the same level since the "I" that appears at the horizon of reflection never appears as the source of the spontaneous consciousness that thinking implies (Sartre 2003b, 124, 127).

Hence it is not clear that the "I" figuring in my articulation of my consciousness (of) thinking also pertains to the structure of my reflecting consciousness, nor that it *already* pertained to the structure of my unreflected consciousness prior to its being reflected. For all we know, it might well be my act of reflection that *endows* my unreflective consciousness (of) thinking with the "I" that I employ to articulate it. We must be all the more alert to this possibility as reflection is apt to *modify* the reflected. Sartre concludes that at the very least reflection cannot itself settle the question whether it confers an "I" or it finds one that was there already.

However, it is obvious that this argument is double-edged. In establishing that the reflective *Cogito* does *not* in fact show that consciousness, *qua* self-conscious, must be centered on an "I", it also establishes by the same token that reflection *cannot* show that consciousness, *qua* self-consciousness, is *not* centered on an "I". It thus cuts both ways. It seems as if Sartre has cut the branch on which he is sitting (see De Coorebyter 2000, 343).

At this point, Sartre invokes the possibility of *unreflective recollection* as a way out of this apparent predicament. Thus, I can unreflectively recollect my past awareness of reading this novel. To do so, it is necessary and sufficient that I should reflect exclusively on the circumstances and the content of my reading (Sartre 2003b, 100-101). On the one hand, it is enough, given that I cannot recall reading this or that in such and such circumstances without thereby resurrecting my past unreflective awareness of doing so. It is not merely that I cannot recall reading without being unreflectively aware of doing so, just as I cannot read without being unreflectively aware of doing so, but rather that I cannot recall reading this or that in such and such circumstances without thereby resurrecting my past unreflective awareness of doing so. On the other hand, provided I reflect exclusively on the circumstances and the content of my reading, in recalling them I can recall my past awareness of reading without reflecting on it and thus without ascribing it. As I direct my recollection on them, I recollect my past awareness of them, but only laterally or indirectly as it were. The past awareness lurks in the background of my present recollection. My present recollection carries it along with itself. I am thus in a position to consult an *unreflected memory* of that awareness. As I consult it, I find that this awareness was

exclusively directed at *what* I was then reading. I was wholly absorbed in my reading. My awareness made no room for an “I” (Sartre 2003b, 101-102).

Thus runs the third and last argument of the *Transcendence of the Ego* against *CE*. The passage in the Introduction to *Being and Nothingness* where the “pre-reflective *Cogito*” makes its first appearance is obviously a descendant of the last cited passage. Recall Sartre’s claim that my spontaneous answer, “I am counting the cigarettes”, “is not only aimed at (*vise*) the instantaneous consciousness which I may reach by reflection, but also at those that have passed without having been reflected, those that are forever *unreflected* in my immediate past” (Sartre 2003a, 9; translation modified). But Sartre no longer holds the certainty of “pure reflection” to be restricted to the instant (see De Coorebyter 2000, 350-352). “I am counting” reflects my counting as something which I *was* already aware (of), it reflects an immediate stretch of past awareness “as being already there” (*comme étant déjà là*) prior to being reflected and so as what makes possible this very reflection (Sartre 2003d, 150).

The reflective consciousness to which “I am counting the cigarettes” gives expression does not endow the unreflected consciousnesses that it reflects with an “I”. In “I am counting the cigarettes”, “I” figures in such a way that it contributes to the disclosure of an “I am thinking” (= *cogito*) in which there is simply no “I” that is thinking—we could call the latter an “I-think” to distinguish it from the “I think” that must be able to accompany any consciousness of mine. The pre-reflective I-think is subjectless.

Once it is revised in accordance with the advances made by *Being and Nothingness*, the third argument put forward by *The Transcendence of the Ego* against *CE*, far from being secondary, proves uniquely compelling, as it brings into display the unity of the other two arguments. It shows the logical impossibility and the logical superfluity of a “self” to be functions of each other. It thus helps us move from the realization that self-consciousness and intentionality are not incompatible to the realization that each is made possible by the other. We begin to understand how self-consciousness, far from disrupting the ontological absorption of all consciousness in its intentional object (see above Section 4.1), can be the principle of such absorption. Which is to say that we begin to understand the sense in which consciousness can be said to be essentially diaphanous.

4.3. The ontological underpinnings of Sartre’s anti-referentialist account of self-consciousness

One should beware of reading the initial account that the Introduction to *Being and Nothingness* offers of the relation between intentionality (consciousness of something) and self-consciousness (consciousness (of) oneself) as self-standing.³⁰ It is perhaps the book’s main claim that the *internal* relation between intentionality and self-consciousness is bound to remain unintelligible as long as it is not analyzed in terms of the categories of being and nothingness. One fateful consequence of a truncated reading of the book is the assimilation—which is nearly as disastrous as it is widespread—of the opposition between thetic and non-thetic consciousness with any or all of the following oppositions: explicit vs. implicit, express vs. tacit, frontal vs. lateral, direct vs. indirect, actual vs. potential, focal vs. peripheral, about *X* vs. concerning *X*, and so on. As will become

clear in the next part (Part 5), behind these faulty assimilations lurks the *Epistemic View of Self-Consciousness*.

We must get to understand the exact content of the suggestion, made in the Introduction of *Being and Nothingness*, that a consciousness's being *non-thetically* conscious (of) itself *is but the other side* of its being (thetically) conscious of its object, while conversely its being *thetically* conscious of its object *is but the other side* of its being (non-thetically) conscious (of) itself.

The suggestion that the following triad of propositions will be understood together or not at all affords a crucial clue here:

- A. One's consciousness transcends itself in the transcendent object at which it is directed, in the sense that it is *nothing but* the positional consciousness of that object.
- B. One's consciousness (of) one's positional consciousness of the object, hence one's consciousness (of) oneself, is *non-positional*.
- C. Self-consciousness, *i.e.* one's consciousness (of) oneself, *isn't about anything*.

Negation essentially occurs in all three propositions in one guise or another. Accordingly, the leading thread of *Being and Nothingness* as a whole is the following claim: the single key to an elucidation of the internal relation between self-consciousness and intentionality lies in a proper account of negation, which account requires in turn a proper understanding of the mutual irreducibility of being (*être*) and nothingness (*néant*). Sartre's main thesis is that the distinction between these two categories is absolute. This means that being (*être*) and nothingness (*néant*) can neither exchange roles nor mix with each other (see Narboux 2015).

To make a long story short, the single key to making all three claims intelligible lies in recognizing that all consciousness is at once, *qua* positional consciousness of the object at which it is directed, an ontological *affirmation* of that object, and, *qua* non-positional consciousness (of) itself, an ontological *negation* (of) itself. More precisely, all consciousness is the ontological affirmation of its object *to the exact extent that* it is the ontological negation (of) itself, and conversely. While in the midst of counting the cigarettes, I am absorbed in counting them in virtue of the fact that my consciousness is negating itself so as to make room for, *i.e.* posit, the cigarettes to be counted. It is one and the same thing for my consciousness to posit some transcendent object as that which it is not and for it to negate itself as not being that transcendent object. For my consciousness to be directed at the cigarettes to be counted just is for my consciousness to negate itself as not being these cigarettes.

Non-positional consciousness (of) oneself is not a lack of positional consciousness. It is the *internal* negation (of) one's positional consciousness of the object. Conversely, one's positional consciousness of the object is the internal negation (of) one's non-positional consciousness (of) oneself, that is to say, the internal negation (of) one's negation (of) oneself. In this sense, all consciousness can be said to be at once anti-symmetric and bipolar.

The reversal of word order ensuing from the ante-position of the reflexive "self" in the term "self-consciousness" perspicuously renders the fact that the intentional relation *to* something and the relation (to) oneself that is but its other side run in opposite directions. My positional consciousness of something and its non-positional consciousness (of) itself are but one and the same bipolar anti-

symmetric relation, $- \rightarrow +$, now read in one direction (from source to target) now read in the opposite direction (from target to source). Self-consciousness is not to be pictured as a second arrow directed at the arrow of intentional consciousness or as an arrow doubling back onto itself, but simply as that arrow itself looked at from its end (see Narboux 2015). This clarifies why the structure exhibited by self-consciousness has nothing to do with that of retroverted intentionality. The notion of *directedness* that figures in our formulation of the *Transparency Principle* (see above 2.1) is misleading. For “world-directedness” is the only directedness there is. Thus we find again that the *Transparency Principle* cannot be accepted as it stands. Again, what it presents as a mystifyingly smooth transition is really no transition at all. Even once it is no longer in service of the “inner gaze model” of self-consciousness, talk of “transparency” can only be provisional by Sartre’s light, given that it maintains in place the picture of consciousness as a medium.

The assimilation of non-thetic consciousness to “implicit consciousness” and of the non-thetic, pre-reflective *Cogito* introduced in *Being and Nothingness* to a “tacit *Cogito*” misses the whole point of Sartre’s opposition between thetic consciousness (intentionality) and non-thetic consciousness (self-consciousness). By misconstruing the opposition between affirmation and negation, it dilutes the opposition between being and nothingness, in contradiction to Sartre’s claim that such dilution is absurd. If Sartre’s analysis is at all on the right track, then the idea that my consciousness *of* something might “concern me” without being “about me” is simply unintelligible. For it is not simply that self-consciousness fails to be retroverted intentionality: retroverted intentionality is *exactly* what self-consciousness *is not*.

5. Conclusion

I have argued in this essay that the *Canonical Equation* of self-consciousness with consciousness of one’s self, together with the *Referentialist View of Self-Consciousness* that underwrites it, must be rejected as unintelligible, on account of the specificity of self-consciousness. The two principles of *Immunity* and *Transparency* have turned out to be unstable compromises, irresolute attempts at accommodating criticisms of the equation within its compass.

But the denial that self-consciousness is consciousness of one’s self is *categorical* in character. In a different terminology, the negative statement, “‘I’ is not a referring expression”, is a *grammatical* one. This means that what it rejects is not an intelligible possibility but a piece of confusion. What is rejected by “‘I’ is not a referring expression” does not make any more sense than what is rejected by “‘I’ does not climb stairs.” There is this difference however: because we are not tempted to give a sense to “‘I’ climbs stairs”, we are not tempted to construe its negation as a privation. By contrast, “‘I’ is not a referring expression” cannot but be heard at first as a privation. Whence the fantasy that “‘I’” could have been or should have been a referring expression? The one thing that I have not done in this essay is to exploit the many resources that the thoughts of Sartre and Wittgenstein afford to address this question.³¹

References

- Anscombe, Elizabeth. 1957. *Intention*. Oxford: Basil Blackwell. New edition, Cambridge (Mass.): Harvard University Press.
- Anscombe, Elizabeth. 1962. "On Sensations of Position." In Anscombe 1981.
- Anscombe, Elizabeth. 1963. "Mental Events." In Anscombe 1981.
- Anscombe, Elizabeth. 1975. "The First Person." In Anscombe 1981.
- Anscombe, Elizabeth. 1965. "The Intentionality of Sensation." In Anscombe 1981.
- Anscombe, Elizabeth. 1976. "The Subjectivity of Sensation". In Anscombe 1981.
- Anscombe, Elizabeth. 1981. *Metaphysics and the Philosophy of Mind, Collected Papers, Volume 3*. Minneapolis: University of Minnesota Press.
- Austin, John L. 1975. *How to Do Things with Words*, second edition, edited by J.O. Urmson and Marina Sbisa. Cambridge: Harvard University Press.
- Balibar, Étienne. 1998. *L'invention de la conscience. John Locke, identité et différence*. Paris: Seuil.
- Balibar, Étienne. 2011. *Citoyen sujet et autres essais d'anthropologie philosophique*. Paris: Presses Universitaires de France.
- Benvéniste, Émile. 1946. "Structures of the Relations of Persons in the Verb". In Benvéniste 1973.
- Benvéniste, Émile. 1956. "The Nature of Pronouns." In Benvéniste 1973.
- Benvéniste, Émile. 1973. *Problems of General Linguistics*, volume 1. Miami: University of Miami Press. First published, in French, in 1966.
- Boyle, Matthew. Unpublished, a. "Transparency and Reflection."
- Boyle, Matthew. Unpublished, b. "Sartre on Non-Positional Bodily Awareness." Paper presented at the conference "Reflection and Non-Positional Self-Consciousness", Patras, Greece, June 19-21, 2017.
- Cassam, Quassim. 1997. *Self and World*. Oxford: Oxford University Press.
- Castañeda, Hector-Neri. 1999. *The Phenomeno-Logic of the I. Essays on Self-Consciousness*. Indianapolis: Indianapolis University Press.
- De Coorebyter, Vincent. 2000. *Sartre face à la phénoménologie*. Brussels: Ousia.
- De Gaynesford, Maximilian. 2017. "Wittgenstein on 'I' and the Self." In *A Companion to Wittgenstein*, edited by Hans-Johann Glock and John Hyman, 478-490. Cambridge: John Wiley & Sons.

- Descombes, Vincent. 2002. "Comment savoir ce que je fais ?" *Philosophie* 76: 15-32.
- Descombes, Vincent. 2004. *Le complément de sujet: enquête sur le fait d'agir de soi-même*. Paris: Gallimard.
- Descombes, Vincent. 2007. "Réponses." In *Vincent Descombes, Questions disputées*. Edited by Bruno Gnassounou and Cyrille Michon. Nantes: Éditions Cécile Defaut.
- Descombes, Vincent. 2009. *Dernières nouvelles du moi*. With Charles Larmore. Paris: Presses Universitaires de France.
- Descombes, Vincent. 2010. "Le marteau, le maillet et le clou." *Revue de métaphysique et de morale* 68, 495-519. Reprinted in Descombes 2014.
- Descombes, Vincent. 2013. *Exercices d'humanité*. Paris: Les petits Platon.
- Descombes, Vincent. 2014. *Le parler de soi*. Paris: Gallimard.
- Descombes, Vincent. 2016. "Response to Richard Moran." Paper presented at the conference "Action and the First Person", Porto, Portugal, June 16-17.
- Evans, Gareth. 1982. *The Varieties of Reference*, edited by John McDowell. Oxford: Clarendon.
- Frank, Manfred. 1991. *Selbstbewusstseinstheorien von Fichte zu Sartre*. Suhrkamp.
- Frank, Manfred. 2007. "Non-Objectal Subjectivity." *Journal of Consciousness Studies* 14, 5-6: 152-173.
- Fried, Michael. 1988. *Absorption and Theatricality: Painting and Beholder in the Age of Diderot*. Chicago: The University of Chicago Press.
- Gardner, Sebastian. 2009. *Sartre's Being and Nothingness: A Reader's Guide*. London: Continuum.
- Geach, Peter. 1972. "Intentionality: On Beliefs about Oneself." In *Logic Matters*. Oxford: Blackwell.
- Gnassounou, Bruno. 2010. "Penser à soi." *Philosophie* 4, 107: 69-92.
- Haddock, Adrian. Unpublished. "I am NN".
- Hagège, Claude. 1974. "Les pronoms logophoriques." *Bulletin de la Société de Linguistique de Paris* 69: 287-310.
- Henrich, Dieter. 1982. "Fichte's Original Insight." *Contemporary German Philosophy* 1: 15-53. First published, in German, in 1966.
- Humbert, Jean. 2004. *Syntaxe grecque*. Paris: Klincksieck. First published in 1945.
- Kaplan, David. 1989. "Demonstratives". In *Themes from David Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein, 481-563. Oxford: Oxford University Press.

Kaplan, David. 2005. "What is Meaning? Explorations in the Theory of Meaning as Use." Unpublished.

Kenny, Anthony. 2001. "Body, Soul and Intellect in Aristotle." In Anthony Kenny, *Essays on the Aristotelian Tradition*, 76-91. Oxford: Oxford University Press. Originally published in James Crabbe, editor, *From Soul to Self* (London: Routledge, 1999).

Kripke, Saul. 2011. "The First Person." In *Philosophical Troubles, Collected Papers, volume 1*, 292-321. Oxford: Oxford University Press.

Larmore, Charles. 2009. *Dernières nouvelles du moi*. With Vincent Descombes. Paris: Presses Universitaires de France.

Lewis, David. 1979. "Attitudes *De Dicto* and Attitudes *De Se*." *The Philosophical Review* 88, 4: 513-543.

Longuenesse, Béatrice. 2006. "Self-Consciousness and Consciousness of One's Own Body: Variations on a Kantian Theme." *Philosophical Topics* 34, 1-2: 283-309.

Longuenesse, Béatrice. 2008. "Self-Consciousness and Self-Reference: Sartre and Wittgenstein." *European Journal of Philosophy* 16, 1: 1-21.

Longuenesse, Béatrice. 2017. *I, Me, Mine*. Oxford: Oxford University Press.

McDowell, John. 2009. "Referring to Oneself." In John McDowell, *The Engaged Intellect: Philosophical Essays*, 186-203. Cambridge (Mass.): Harvard University Press. Originally published in Lewis E. Hahn, editor, *The Philosophy of P.F. Strawson* (Chicago: Open Court, 1998).

McDowell, John. 2011. "Anscombe on Bodily Self-Knowledge". In *Essays on Anscombe's Intention*, edited by Anton Ford, Jennifer Hornsby, and Frederick Stoutland, 128-146. Cambridge (Mass.): Harvard University Press.

McDowell, John. 2012. "Réponse à Valérie Aucouturier." In *Autour de L'esprit et le monde de John McDowell*, edited by Anne Le Goff and Christophe Al-Saleh, 280-284. Paris: Vrin.

McDowell, John. Unpublished. "Some questions for Matthew Boyle." Paper presented at the conference "Reflection and Non-Positional Self-Consciousness", Patras, Greece, June 19-21, 2017.

Miguens, Sofia; Morando, Clara Bravo; Preyer, Gerhard, editors. 2016. *Pre-reflective Consciousness: Sartre and Contemporary Philosophy of Mind*. New York: Routledge.

Moran, Richard. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.

Narboux, Jean-Philippe. 2015. "Intentionnalité et négation dans *L'être et le néant*," 57-91. In *Sartre. L'être et le néant. Nouvelles lectures*, edited by Jean-Marc Mouillie and Jean-Philippe Narboux. Paris: Les Belles Lettres.

Narboux, Jean-Philippe. Forthcoming 2017a. "Pensées en première personne et cogitationes cartésiennes." In *Les formes historiques du Cogito*, edited by Kim Sang Ong-Van-Cung. Paris: Classiques Garnier.

Narboux, Jean-Philippe. Forthcoming 2017b. "Speaking of Oneself and Speaking of One's Self." Review essay on *Le parler de soi*, by Vincent Descombes. *European Journal of Philosophy*.

Peacocke, Christopher. 2014. *The Mirror of the World: Subjects, Consciousness and Self-Consciousness*. Oxford: Oxford University Press.

Pippin, Robert. 2014. "The Significance of Self-Consciousness in Idealist Theories of Logic." *Proceedings of the Aristotelian Society* 114, 2: 145-166

Rödl, Sebastian. 2007. *Self-Consciousness*. Cambridge (Mass.): Harvard University Press.

Rödl, Sebastian. 2016. "The First Person and Self-Knowledge in Analytic Philosophy." In *Self-Knowledge: a History*, edited by Ursula Renz, 280-294. Oxford: Oxford University Press.

Sartre, Jean-Paul. 2002. "Intentionality: a Fundamental Idea of Husserl's Phenomenology", translated by Joseph Fell. In *The Phenomenology Reader*, edited by Dermot Moran and Timothy Mooney. London: Routledge. First published, in French, in 1939.

Sartre, Jean-Paul. 1980. *L'être et le néant: essai d'ontologie phénoménologique*. Paris: Gallimard. First published in 1943.

Sartre, Jean-Paul. 2003a. *Being and Nothingness*, translated by Hazel Barnes, introduction by Mary Warnock with a new preface by Richard Eyre. London: Routledge. First published, in French, in 1943.

Sartre, Jean-Paul. 2003b. *La transcendance de l'ego*. In *La transcendance de l'ego et autres textes phénoménologiques*, edited by Vincent de Coorebyter, 93-131. Paris: Vrin. First published in 1937.

Sartre, Jean-Paul. 2003c. "Une idée fondamentale de la phénoménologie de Husserl." In *La transcendance de l'ego et autres textes phénoménologiques*, edited by Vincent de Coorebyter, 87-89. Paris: Vrin. First published in 1939.

Sartre, Jean-Paul. 2003d. "Conscience de soi et connaissance de soi." In *La transcendance de l'ego et autres textes phénoménologiques*, edited by Vincent de Coorebyter, 135-165. Paris: Vrin. First published in 1947.

Sartre, Jean-Paul. 2004. *The Transcendence of the Ego*. London: Routledge. First published, in French, in 1937.

Sartre, Jean-Paul. 2016. *What is Subjectivity?*, translated by Trista Selous, introduction by Michel Kail and Raoul Kirchmayr, afterword by Fredric Jameson. London: Verso. First published, in French, in 2013.

Shoemaker, Sidney. 1968. "Self-Reference and Self-Awareness," *The Journal of Philosophy* 65, 19: 555-567.

Strawson, Peter. 1959. *The Bounds of Sense*. London: Methuen.

Strawson, Peter. 1966. *Individuals*. London: Methuen

- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.
- Thompson, Michael. 2008. *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge (Mass.): Harvard University Press.
- Thompson, Michael. 2011. "Anscombe's *Intention* and Practical Knowledge." In *Essays on Anscombe's Intention*, edited by Anton Ford, Jennifer Hornsby, and Frederick Stoutland, 198-210. Cambridge (Mass.): Harvard University Press.
- Thompson, Michael. Unpublished. "You and I: Some Puzzles about 'Mutual Recognition'."
- Tugendhat, Ernst. 1986. *Self-Consciousness and Self-Determination*, translated by Paul Stern. Cambridge (Mass.): The MIT Press. First published, in German, in 1979.
- Webb, Samuel. 2016. "Savoir ce que je fais: Anscombe et Sartre. Vers une étude comparative". *Klesis*, 35: 12-30.
- Wittgenstein, Ludwig. 1953. *Philosophical Investigations*, translated by Elisabeth Anscombe.
- Wittgenstein, Ludwig. 1958. *The Blue and the Brown Books*.
- Wittgenstein, Ludwig. 1980. *Remarks on the Philosophy of Psychology, volume 1*. Cambridge: Basil Blackwell.
- Zahavi, Dan. 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge (Mass.): MIT Press.

Notes

¹ In a fuller version of the present essay, I argue that *CE* also conceals another and equivalent substantial claim, namely that "Self-Consciousness is self-knowledge" (*Epistemic View of Self-Consciousness*). Accordingly, I maintain against *RSC* and *ESC*:

N_R "I", *qua* manifestation of self-consciousness, is *not* a referring expression.

N_E "I", *qua* manifestation of self-consciousness, is *not* a manifestation of knowledge.

I try to explain why Sartre and Wittgenstein both resist the assimilation of self-consciousness to self-knowledge.

² A notable exception is Béatrice Longuenesse's recent work on the first person, to which this essay as a whole is much indebted, however divergent its conclusions may be from hers. See especially Longuenesse 2008 and Longuenesse 2017.

³ One simple way of formulating this fundamental divide is as a divide over the legacy of Descartes's *Cogito*. At the acme of the *First Meditation*, Descartes raises the following question for himself: "Someone is thinking the thought that 'I am being fooled by the evil demon', but is it I?" According to Descartes, in realizing for myself that I cannot make sense of this doubt, I recognize these three truths:

C_1 For me to be conscious is for me to be conscious of being conscious, self-conscious. (Existence of self-consciousness)

-
- C_2 For me to be conscious of being conscious, self-conscious, is for me to be conscious of the being of my self. (Existence of “I”)
- C_3 For me to be conscious of being conscious, self-conscious, is for me to be conscious of my self as being a conscious thing. (Essence of “I”)

Virtually all of Descartes’s readers, including those who belong to the two aforementioned schools of philosophy, have granted the first implication, thereby embracing *CE*, and denied the second. By contrast, Sartre and Wittgenstein (and Anscombe after him) have denied the first implication, thereby rejecting *CE*, and granted the second (see Descombes 2013; Narboux 2017a). This accounts for what would be otherwise a baffling feature of the dispute between proponents and detractors of the *Referentialist View of Self-Consciousness*, namely that each side charges the other side with relapsing into a form of Cartesianism that is deeper and more gripping than the form it professes to overcome. Because Sartre and Wittgenstein are not concerned to criticize the second implication (Anscombe goes so far as to concede the conditional truth of C_3), they have been said to be in the grip of some “ultimate Cartesian error” (Strawson 1959; McDowell 2009). They have been charged with replacing Descartes’s dualism between two radically distinct kinds of reference with a dualism of reference and non-reference or, as one might also put it, with a dualism of being and nothingness (see Strawson 1959, 98). From the standpoint of Sartre and Wittgenstein, however, mainstream phenomenology and mainstream analytic philosophy have erred alike in failing to take the criticism of Descartes’s *Cogito* far enough. They have denounced the Cartesian specification of the reference of “I”, when what needed to be questioned was the very assumption that “I” so much as referred to anything in the first place.

⁴This recalls a famous passage in Sartre’s article, “Intentionality: a Fundamental Idea of Husserl’s Phenomenology”: “There is nothing in [consciousness] but a movement of fleeing itself, a sliding out of itself. If, *per impossibile*, you could enter ‘into’ a consciousness, you would be seized by a whirlwind and thrown back outside, near the tree, in the thick of the dust, for consciousness has no ‘inside’.” (Sartre 2002; translation modified)

⁵Wittgenstein writes: “There are two different cases in the use of the word “I” (or “my”) which I might call “the use as object” and “the use as subject”. Examples of the first kind of use are these: “My arm is broken”, “I have grown six inches”, “I have a bump on my forehead”, “The wind blows my hair about”. Examples of the second kind are: “I see so-and-so”, “I hear so-and-so”, “I try to lift my arm”, “I think it will rain”, “I have toothache”. One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these cases the possibility of an error, or as I should rather put it: The possibility of an error has been provided for. (...) It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbour’s. And I could, looking into a mirror, mistake a bump on his forehead for one on mine. On the other hand, there is no question of recognizing a person when I say I have toothache. To ask “Are you sure that it’s *you* who have pains?” would be nonsensical.” (Wittgenstein 1958, 66-67)

⁶First, we are endowed with what Evans characterizes as “a general capacity to perceive our own bodies”, a capacity comprising “our proprioceptive sense, our sense of balance, of heat and cold, and of pressure” (Evans, 1982, 220). Second, “we are able to know our position, orientation, and relation to other objects in the world upon the basis of our perceptions of the world” (Evans 1982, 222): in short, we can locate ourselves relative to objects of outer perception upon the basis of our perceptions of them (see below Section 6.2) Third, we can have knowledge of our own intentional bodily movements and more generally of our own intentional actions, as we ourselves move and act in the world (Evans 1982, 207, 224 note 34; McDowell 2009).

Capacities of all three types give rise to judgments that are immune to error through misidentification. When the relevant capacities are actualized, it does not make sense to ask “Someone’s legs are crossed, but is it my legs that are crossed?”, “Someone is hot and sticky, but is it I who I am hot and sticky?”, “Someone is being pushed, but is it I who am being pushed?” (Evans 1982, 216, 221) Neither does it make sense to ask “Someone is in my bedroom, but is it I?”, “Someone is moving, but is it I?”, “Someone is standing in front of a tree, but is it I?” (Evans 1982, 222) Nor, finally, does it make sense to ask “Someone is pushing the boulder, but is it I?”

In other words, in none of these three cases is it possible to drive a wedge between one's awareness (or apparent awareness) that a certain physical property (say, the property of being F) is being instantiated and one's awareness (or apparent awareness) that it is being instantiated by oneself. In all such cases, being aware (or apparently aware) that a certain physical property holds of someone just is being aware (or apparently aware) that it holds of oneself.

⁷ First, Evans's characterization of the mediating component runs afoul of Anscombe's denial that statements of the form < I am b > are identity statements, which is an immediate consequence of her denial that "I" is a referring expression (Anscombe 1975, 33; McDowell 2009, 195). Second, in Anscombe's view, it is at best very misleading to trace the unmediated knowledge that we can have of our own bodily postures and motions to "a faculty of internal self-scanning" (Evans 1982, 230), thereby entertaining the false suggestion that 'sensations' of position stand on par with sensations in the proper sense and only differ from the latter in that they are turned inwards (see Anscombe 1962, 72-73).

⁸ This suggests that there is more than one way of understanding the diagnosis made in this passage: "Our thoughts about ourselves are in no way hospitable to Cartesianism. Our customary use of 'I' simply spans the gap between the mental and the physical, and is no more intimately connected with one aspect of our self-conception than the other." (Evans 1982, 256) Although there is a sense in which Sartre and Anscombe wholeheartedly endorse this diagnosis, on a second look, the agreement turns out to be little more than verbal. Evans's own way of fleshing out this diagnosis turns on the invocation of what he calls "the Generality Constraint" (see below Section 6.2). The crucial point, for Evans, is not so much that *Immunity* is not confined to the realm of narrowly psychological self-ascriptions, as that it enables self-thinking to meet the Generality Constraint by allowing self-identification with a bit of matter to get a grip. But on the very different view that Sartre and Anscombe take of the above diagnosis, Evans's idea evinces a relapse into Cartesianism. Thus we find, unsurprisingly, that the dispute over the legacy of Wittgenstein's reflections on "I" broached in the present section is echoed in the dispute over the legacy of Descartes's *Cogito* broached in Part 1.

⁹ Even though an I-statement does not by itself tell its addressee of whom (*i.e.* which object) it speaks, since the I-sentence asserted by X mentions no ordinary proper name of X, it does not express a general thought. Unlike the truth-conditions of a general statement (like "Somebody shall be banished" or "The murderer of Laius shall be banished"), and like those of a singular statement (like "This man over there shall be banished"), the truth-conditions of an I-statement cannot be grasped apart from some knowledge of whom (*i.e.* which object) it speaks (Descombes 2014, 289).

¹⁰ In order to expose this mistake, it will be expedient to imagine, after Anscombe, a society in which everyone is labeled by two names, the second of which (say, "Alex") is stamped on everyone's wrist (in that way, it is the same for everyone), but is used by each whenever he is to make reports on his own actions (in that way, it is used by each only to speak of himself)—which he does only on the basis of observation, or at any rate never *without* observation (see Anscombe 1975, 24-25). The name "Alex" obviously conforms to the same rule as "I": "Alex" designates the one who uses it. If we call a report whose syntactical subject is "Alex" an "Alex-report", then we have the following truth-functional equivalence: an Alex-report issuing from the mouth of an individual named "N" is true if and only if what it reports is true of N. Like I-reports, Alex-reports do not by themselves tell their addressee of whom they each speak. For all that, their truth-conditions are not general but singular. Again, Alex-reports are just like I-reports in this respect. Any Alex-report issuing from the mouth of N will be about the object that N (in fact) is, since it will report an action of the object that N (in fact) is. For N to make an Alex-report is for N to speak of himself. Any observation-based Alex-report issuing from the mouth of N will rest on N's observation of an action of the object that N (in fact) is. Therefore, any such report will be a function of N's consciousness of himself. And yet it is obvious that one's consciousness of oneself, in that sense, falls short of what we mean by "self-consciousness", and which manifests itself in our use of "I". Thus, the difference between "I" and an ordinary proper name is not a trivial one. It is far from reducing to the fact that everyone makes use of "I" only to speak of himself. That the members of the imagined society are competent Alex-users goes no way towards establishing that they do not lack self-consciousness (see Anscombe 1975, 24-25, 36).

¹¹ On this dialectic, see Haddock, unpublished.

¹² Consider, for example, these three pairs of statements: (a) “Through that single slip of the tongue, the man revealed that he_D was the guilty part”, (b) “The man confessed that he_I was the guilty part”; (a’) “The prisoner harmed himself_D over the barbwire while running for escape” (b’) “The prisoner admitted to have harmed himself_I so as to be taken to the infirmary”; (a’’) “In declaring that the murderer of Laius was to be banished, Oedipus, unbeknownst to himself, declared of himself_D that he_D was to be banished”, (b’’) “In declaring that the murderer of Laius was to be banished, Oedipus did not declare that he-himself_I was to be banished”. In all three cases, the content reported in the second statement cannot be specified otherwise than by a counterpart first-person utterance placed in the mouth of the individual whose operation is reported (Castañeda 1999; Geach 1974, 128-129; Anscombe 1975, 22): what the man confessed is: “I am the guilty part”, what the prisoner admitted is: “I have harmed myself so as to be taken to the infirmary”, what Oedipus failed to declare is: “I am to be banished”.

In English, as indeed in most languages, the distinction between the direct and the indirect reflexive is not marked morphologically. This makes for some ambiguity. Thus the utterance “Out of anger, the child locked herself in her room” admits of two incompatible understandings, according to whether the feeling of anger is apprehended as the cause of a slip (in which case “herself” must be construed as the direct reflexive) or as the reason of an intentional action (in which case “herself” must be construed as the indirect reflexive). Castañeda famously resorted to the star-symbol “*” to remedy this defect, using it as a marker of the indirect reflexive (Castañeda 1999). In Attic ancient Greek, there exists a special form for the indirect reflexive whose use is in order whenever the completive clause in which the reflexive occurs reports the thoughts or words of the syntactical subject of the main clause (Humbert 2004, 63; Anscombe 1975, 22; Descombes 2004, 138-139). This is also true of some West African languages. Claude Hagège coined the term “logocentric pronoun” to designate the special form present in these languages (see Hagège 1974; Thompson, unpublished).

Thus, the indirect reflexive can be seen as a device for mapping a first-person statement onto a third-person statement (Descombes 2004, 138; Descombes 2010, 503). The crucial point is that it enables language to cast a first-person statement in the mould of a third-person statement without, for all that, cancelling the irreducibility of the former to the latter. The availability of such a device makes it intelligible that a third-person statement could be expressive of self-consciousness and even express an immediate thought, despite the logical originality of the first person and the asymmetry between the first and third persons of psychological verbs in the present indicative.

¹³ Saying that “I” has the *syntactical* look of a proper name (Anscombe 1975, 23; Descombes 2009, 92) is not the same as saying (as John McDowell puts it in a passage intended as a reformulation of Anscombe’s view) that “I-statements have, as it were, the look of predications, with “I” in subject position” (McDowell 2009, 187) (where the notion of predication is a *logical* one), let alone the same as saying that I-statements look *to be* predications. As Descombes underscores, “Anscombe nowhere suggests that the ordinary use of first-person forms of expression is a source of mystification” (Descombes 2010, 504) and for her “the myth of the *self* does not originate in ordinary language.” (Descombes 2010, 511) It can be traced to the grip exerted by the phrase “self-consciousness”, but this phrase originates in philosophy, as Anscombe notes (Anscombe 1975, 25).

¹⁴ Accordingly, the true difference between “I” and our invented proper name “Alex” is that the use of “Alex” does involve the apprehension of the referents of its repeated occurrences in the mouth of one and the same speaker as being the same (Anscombe 1975, 27). “Alex” is so defined that it cannot fail to get hold of an existing object, but it can fail to get hold of the right object. In effect, the “Alex”-user might mistake someone else than Alex for “Alex” (Anscombe 1975, 30). If “I” were a regular proper name, then it would be liable to that kind of mistake.

¹⁵ It is worth reviewing some of these features, if only to dissipate the misunderstandings that attach to them. First, that the utterer of “I”, call her “N”, does not thereby identify herself, obviously does not mean that her identity (her being N) is immaterial to the truth-conditions of the predicative utterance to which her first-person utterance is truth-functionally equivalent in virtue of the rule governing the use of “I”. If N tells me “I am reading *Ulysses*”, not only am I entitled to hold N to be telling me that she* is reading *Ulysses*, but also what I am being told by N is true if and only if N is in fact reading *Ulysses*. What is more, barring a few odd cases, when N

tells me “I am reading *Ulysses*”, her utterance enables me (and is put forward by her so as to enable me) to *know who* (i.e. *which* person) is being said to be reading *Ulysses* and in that sense to *identify who* is being said to be doing so. I need only advert to the mouth out of which, or to the grain of the voice with which, “I” is being proffered. *If* moreover I happen to know that the utterer of “I” is called “N” (in contrast with “M”, “O”, “P”, etc.), then her utterance even enables me to *identify* the person who is being said to be reading *Ulysses as being N* (rather than M, O, P, etc.).

We remarked early on, in Section 2.3, that what shows forth in my I-sentences, is not *who* I am speaking of, but only *that* I am speaking of myself. The other side of the coin is that the task of identifying the speaker is assigned by the speech-act to its *addressee* (Descombes 2004; Descombes 2014). Far from expressing an act of self-identification prior to the speech-act, “I”, like all grammatical persons, serves to identify an individual *only in virtue of* the position that the latter occupies in the speech-act. The function of identification accruing to “I” is *mediated* by the actual speech-act. Thus, it is not enough to say that “I” designates the one uttering it. It would be far more accurate to say that “I” designates the one *among the two of us* currently addressing himself or herself to the other, in other words, the one presently performing a bipolar speech-act, i.e. an I/You-speech-act (Descombes 2004; Descombes 2007, 403). The above qualification—the exclusion of “odd cases”—is required because we need to make room for the (logical) possibility that the body out of which “I” is proffered (say, N’s body) should be “possessed” by someone else (say, by M) (Wittgenstein 1957; Anscombe 1975, 33).

Second, that someone called “N” does not identify herself when she speaks of herself in the first person of course does not imply that she cannot identify herself, since she may do so by speaking of herself in the third person, if only by using her own name (see above Section 2.3). When N speaks of herself in that way, say by uttering the sentence “N was born in London”, her utterance will encapsulate or even constitute itself an act of “self-identification”, in what are two perfectly acceptable senses of that phrase. First, her utterance of the sentence “N was born in London” presupposes an act of “self-identification” to the extent that the name “N” singles her out as the reference of the logical subject to which “was born in London” is predicated. If moreover in the context at hand the utterance “N was born in London” counts as an answer to the query “who (among M, N, O, P) was born in London?” rather than as an answer to an alternative query like “in which town was N born?” then N herself can be said to be singling herself out by means of these words, and to be identifying herself in the sense of “casting” herself in the role of one born in London.

Third, finally, that the utterer of “I” does not thereby identify herself, certainly does not imply that for someone to declare “I am N” is not for her to convey information about her own identity. Declaring “I am N” is tantamount to presenting a piece of ID. In declaring “I am N”, I may either identify (= “cast”) myself as being N or identify (= “place”) my name as being N, as the case may be, according to whether my utterance qualifies as an answer to the question “which among these names (“M”, “N”, “O”, “P”) is yours?” or to the question “who among you is called N?” In order to convey information about the speaker’s identity, the proposition “I am N” need not be an identity proposition. In any case, the *loss of identity* that can result in the inability of someone named N to recognize that she* is N (where “she” is the indirect reflexive), or that *that* person whom she would point to in pointing to herself (where “herself” is the direct reflexive) is N, has nothing to do with the *lapse of self-consciousness* evinced by Baldy in James’s Story (see above Section 2.3) (Anscombe 1975, 33; Larmore 2009, 122)

¹⁶ See Benvéniste 1946, 230-231; Benvéniste 1956, 255; Tesnière 1959, 117.

¹⁷ Anscombe makes the same point about a related case in this passage: “If something is a correct observation, we ought to be able to detach what is observed from the statement of the observer’s perception of it, and simply say it was so. / Now it is perfectly true to say that the explicit statement of the case is “I see these reds to match”, and if “I” is a name of something involved in this, it is difficult to see how the detachment could ever be justified.” (Anscombe 1976, 54)

¹⁸ Anscombe writes: “It may be senseless to speak of my having reasons, grounds, for what I say, or it may be good sense. It is senseless in the case of “I have a headache” and good sense in the case of “I believe his story”. Senseless in the case of “I have a headache”, for what could the ground be? The headache? The datum itself? How can a sensation imply anything? It would have to be the statement of sensation that was the ground: but that is simply “I have a headache”, which was the statement in question. It is good sense in the case of “I believe his story”: but then

one's grounds are grounds for believing, i.e. grounds tending to show that the thing believed is true. In either case, there is a *dissimilarity* between the first and the third person in the simple (assertoric) present indicative: consider the grounds one might have for "He's got a headache" and "He believes that story". Equally, though I may have grounds when I say "I want an apple", these would be, generally, not reasons for *judging* that I want an apple, but reasons for wanting an apple, whereas grounds for "He wants an apple" should be reasons for *judging* that he wants an apple." (Anscombe 1963, 60-61)

¹⁹ In *CE*, the English term "self" translates as before the terms "*moi*" and "*ego*". Sartre draws a sharp distinction between the "Ego" (which I translate as "the self", following the convention inaugurated by Locke) and "*le Soi*" (which I translate as "the Itself"), and "*le soi*" (which I translate as "the ipse") just as he draws a sharp distinction between two concepts of reflection, namely "reflection" in the sense of "*réfléchir*" and "reflection" in the sense of "*réfléter*". I reserve the term "reflection" for the former. An examination of the significance of these two distinctions falls outside the purview of this essay. Suffice to say that many misunderstandings of Sartre turn on a lack of attention to these distinctions.

²⁰ What makes it possible to bring together Sartre's concept of intentionality and Anscombe's concept of reference is that Sartre, unlike Anscombe (see Anscombe 1965, 4), conceives intentionality as an object-dependent relation and not as a purely internal relation, and that Anscombe underlines that the relation of reference at stake in her discussion is object-dependent.

²¹ My interpretation of the mutual relevance of Sartre and Anscombe's thoughts thus departs from the ones propounded by Vincent Descombes and Béatrice Longuenesse. Although they assess Anscombe's diagnosis in opposite ways (Descombes endorses it whereas Longuenesse rejects it) and although they read in opposite ways Sartre's evolution between *The Transcendence of the Ego* and *Being and Nothingness* (Descombes regards Sartre's mature account as more consistently paradoxical than his earlier account whereas Longuenesse maintains in her 2008 essay that the deflationary account of self-consciousness advocated by Sartre in his early work is importantly qualified and mitigated in his later work), they concur on the view that Sartre's mature account is irremediably paradoxical. In brief, Longuenesse reads Anscombe in the way that they both read Sartre. Descombes reads Sartre and Anscombe as lying at the opposite ends of the philosophical spectrum.

According to Descombes, Sartre does not challenge *RSC* so much as he brings it to a climax (Descombes 2009, 78). Descombes concedes that Sartre is fully cognizant of the paradoxes attending any account of self-consciousness in terms of self-reference (Descombes 2004, 106; Descombes 2009, 78-79). But he maintains that Sartre, for lack of a better alternative, is unable to forsake this model, so that he ends up reading these paradoxes into the structure of self-consciousness itself (Descombes 2004, 106; Descombes 2009, 79). To be more accurate, on Descombes's reading, Sartre wavers between tracing these paradoxes to the very structure of self-consciousness and trying to lift the air of paradox by helping himself to typographical distinctions—such as the one between two notions of "of-ness", respectively marked by the lack or presence of parentheses around the preposition. In the one case, Sartre admits to finding no satisfactory—because no non-contradictory—formulation for his views and he indulges in "dialectical formulas" (Descombes 2004, 106-107; Descombes 2009, 79). In the other case, Sartre contrives a rhetorical expedient in order to keep the contradiction at bay (Descombes 2016).

I take issue with both horns of this dilemma in this part. I address elsewhere Descombes's objections against Sartre's analysis of the syntactical role played by the third-person reflexive ("*soi*", "*se*") in pronominal constructions like "*Il s'ennuie*" ("He is getting bored") or "*Il se penche*" ("He is leaning") at the beginning of his inquiry into the ontological structure of non-thetic consciousness (of) oneself at the beginning of Part 2 of *Being and Nothingness* (Descombes 2004, 105-107; Descombes 2009, 78-79). See Narboux 2017b.

²² On this score, the present reading is in disagreement with the one propounded in Longuenesse 2008.

²³ It is reminiscent of Fichte's negative insight into the insuperable paradoxes plaguing the proposal that self-consciousness is the act in which the self turns back into itself (see Henrich 1982; Frank 1991; Frank 2007). Unlike Fichte, however, and like Wittgenstein and

Anscombe (see Tugendhat 1986; Descombes 2004), Sartre does not think that the alternative proposal according to which “The self posits itself as self-positing” is a satisfactory way out of these paradoxes.

²⁴ On this convention, “self-consciousness” is *by definition* “consciousness (of) itself”, which is necessarily consciousness of itself*/oneself*. Had Sartre been writing in English, he could have dispensed with the device of the parentheses around the preposition “of” and instead resorted to that alternative notational expedient. The two notations are strictly equivalent with respect to both expressive power and perspicuity.

²⁵ See e.g. Sartre 2003d, 150: “I’m reading [*Je suis en train de lire*]. I reply to you, ‘I’m reading’ [*Je lis*], as you ask me what I am doing.”

²⁶ On the significant convergences between Sartre and Anscombe on the topic of intentional action, see Webb 2016.

²⁷ Certainly, in calling unmediated agent-or-patient conceptions “reflective consciousnesses of states, actions, motions, etc.” (Anscombe 1975, 34), Anscombe does not mean that they rest on reflection (*réflexion*) in Sartre’s sense (see Longuenesse 2017, 69-70).

²⁸ At any rate, giving such answers can very well be “spontaneous” rather than “deliberate” (on this contrast, see Anscombe 1957, §27, 47). When she motivates her methodological shift to an analysis of the concept of intentional action by noting that we can safely identify at least some of a person’s intentions simply by saying “straight off” what she actually did or is doing, Anscombe says that most of the things that one would thus say “would coincide with what [s]he could say [s]he was doing, perhaps even without reflection, certainly without adverting to observation.” (Anscombe 1957, §4, 8)

²⁹ Kant’s dictum, with its notable emphasis on the words “be able”, appears to imply: first, that “I think:” (or “I think that”) is *entitled* to be prefixed to the expression of any thought whatsoever (a characteristically *de jure* claim); second, that as a matter of fact “I think” *need not* be so prefixed, that it may well “part company” with a given thought (a *de facto* claim).

This still leaves room for understanding in a variety of ways the claim that the “potential for reflectiveness” (McDowell 2009, 201) that essentially pertains to our thoughts need not be actualized in a given case.

Not only is Kant’s position compatible, for all he says, with “I think” never actually occurring, but also, as a matter of fact, he appears to make this very point. In any case, given the nature of his inquiry, there is no need or him to put forward any particular view about the nature of the relation that someone’s *empirical* uses of “I”—as distinct from the use of “I” in the logical prefix “I think”—bear to the human being that she is (McDowell 2009, 194-195).

³⁰ Readers of Sartre’s treatise have all too often shown prone to confine themselves to the conceptual resources of its introduction in articulating his view of this relation. However, the Introduction makes no claim to have made this relation intelligible, quite the contrary. As one would expect from an introduction, it does not solve a problem, so much as it delineates one. Only by the end of the first half (i.e. first two parts) of the book has the relation in question become intelligible. The progression of the book cannot be understood apart from the *regressive method* underwriting that progression.

³¹ Earlier versions of this essay were presented at the University of Porto (2012), the University of East Anglia in Norwich (2013), Auburn University (2013), and Patras University (2017). I am grateful to the members of these audiences for insightful comments and objections. I am especially indebted to discussions with Arata Hamawaki, Eric Marcus, Matthew Boyle, Adrian Haddock, James Conant, Oskari Kuusela, John McDowell, Sofia Miguens, Richard Moran, Sebastian Rödl, and Charles Travis.

This essay was completed during a stay at the Institute for Advanced Study of Berlin (*Wissenschaftskolleg zu Berlin*) in 2017-2018. During this period, I benefited from a Prestige Co-Financing Grant Award (Marie Skłodowska Curie Actions COFUND) and from an International Mobility Grant by the Career Development Program of the University of Bordeaux (IDEX). I am grateful to these three institutions for their generous support. Finally, I am most grateful to Manu Goswami for her precious help with the ultimate draft.