Varieties of Internal Relations: Intention, Expression and Norms*

Josep Lluís Prades

RESUMEN

En sus ataques a los análisis neo-kripkeanos, McDowell ha aceptado el supuesto de que las atribuciones de intención son normativas en el mismo sentido en que lo son las atribuciones de significado. Propondré que tal asimilación no es correcta. Mencionaré algunas ideas de Wittgenstein sobre la intencionalidad (alrededor de 1930) que habían de preservarse en las *Investigaciones filosóficas*. Trataré de rastrear un argumento del que puede concluirse la conducta expresiva como el proto-fenómeno de la intencionalidad. Las características de tal noción permiten justificar las ideas de McDowell sobre la imposibilidad de fundamentar el "lecho rocoso" de las convenciones gramaticales. Sin embargo, las razones últimas para tal imposibilidad son ligeramente diferentes de las que ha defendido McDowell.

ABSTRACT

In his attack on neo-Kripkean accounts, McDowell has accepted that attributions of intention are normative, in the same sense in which attributions of meaning are normative. I will argue that this is a wrong assimilation. By referring to certain of Wittgenstein's ideas on intentionality (*circa* 1930) that were preserved in *Philosophical Investigations*, I will try to track an argument from which it follows that expressive behaviour is the proto-phenomenon of intentionality. The features of this notion justify McDowell's ideas about the impossibility of grounding the "bedrock" of grammatical conventions. Nevertheless, the underlying reasons for such impossibility are slightly different from those that McDowell has defended.

Wittgenstein was always obsessed by the question of the logical must. In *Philosophical Investigations*, the hardness of the logical must is linked to certain internal connections that are constitutive of intentional states: the connection between a desire, a belief or an intention and their intentional objects. Professor McDowell has consistently attacked those interpretations of Wittgenstein's thinking that devaluate these internal connections and the role they play in his reflections on following a rule. My overall agreement with McDowell is compatible with the main purpose of this paper: I will try to show certain differences between the kind of internal relations that are proper of intentionality, in general, and the more specific normativity that must be

137

involved in following rules. By adopting a slightly different perspective from McDowell's, my hope is to defend his main conclusions but to show, nevertheless, that certain conditions for the possibility of language are not given proper consideration in his account. If I am right, some of the foundations that McDowell requires for his main conclusions must be located in different, albeit very close, places.

I

To begin, let us revisit Kripke's well known challenge to the idea that a particular intention of a person might fix her meaning a determinate mathematical function [Kripke (1982)]. His main point is against the dialectical import of such a strategy. Intention is no less normative than meaning: no fact of the matter seems to determine what the content of the putative intention is. Certain steps in Kripke's argument may be naturally interpreted as grounding a deflationary, "no-fact-of-the-matter" account, not only of meaning but of intention itself. McDowell would disagree with this consequence of Kripke's argument, and so would I. Kripke was searching in the wrong direction. A mental process of forming an intention can determine certain content simply because it has some links with some processes that Kripke did not consider: processes in virtue of which certain mental happenings, with a determinate content, are possible. I am not interested, now, in the particular way in which this rejection of a Kripkean conclusion can be justified, or even in discussing the particular role that it must play in Kripke's overall sceptical argument. I am interested in McDowell: he obviously thinks that there is a way of blocking a Kripkean version of the "no-fact-of-the-matter" argument in the case of intentional content. I agree with him. However, the question in which I am interested is the following: how is this particular way of blocking the standard "no-fact-of-the-matter" argument for intentional content connected to McDowell's own description of the conditions of possibility of following a linguistic rule? For instance, when arguing for the conclusion that following a rule requires a communal practice, McDowell insists that this requirement has to be derived from the rejection of the idea that grasping meaning is always a case of interpretation. So, we have two options: (a) in the general case of intention, we can derive a symmetrical requirement from the rejection of the idea that intentional content requires interpretation; or (b) we must justify why, in the case of following a linguistic rule, the community plays a crucial role, and a crucial role *simply because* meaning and understanding cannot be assimilated to interpretation. I do not see that we can argue for (a): an implicit reference to a communal practice is not in any relevant sense needed to ground the attribution of a basic intentional content to a baby or to an isolated animal. When an adult human being, who is a competent speaker, has an in-

tention, he does not always in any relevant sense *ipso facto* have the intention to adjust his behaviour to a communal practice. We need then to defend (b). Nevertheless, if such a defence is possible — and I will attempt this later — it will have to use, as a crucial premise, not only the general rejection of a mythological notion of interpretation, but something much more specific: the specific content of the propositional attitudes that should be involved in anything worthy of the label "following a linguistic rule".

II

Before proceeding to the argument, let us make a diagnosis about the philosophers' difficulties regarding the determination of intentional content, once they have seen that this content cannot be fixed by a process of interpretation of detached, internal objects. I agree with McDowell that the crucial fact is that this content is determined in virtue of certain links with phenomena outside those putative detached internal objects. Particular thoughts are not detached particulars deprived of intentionality to which an interpretation has to be added. Intention is not a kind of flag in the mind that both points on its own in a determinate direction and is fully independent of, for instance, certain dispositions of the living animal that I am. The crucial point is this: there is no way of describing these connections if this description has to be made in non-intentional language. This is the reason why there can be no satisfactory answer to any question of the form: "in virtue of what fact does a certain event E have the intentional content IC?" The only possible *reaction* is to show that the question itself depends on a crucial misunderstanding. If the fact of the matter is stated by using intentional language, the philosopher who has been attracted by the question will not be happy: he would consider it as a new statement of the fact that is supposed to be analysed. But it cannot be answered in any other language. This is, I would guess, completely symmetrical to certain confusions regarding the fact that a given normative practice of following a rule is internally related to a certain set of actual and possible applications. And the diagnosis regarding the confusions of levels is symmetrical to McDowell's well known insistence that we should not try to dig below the bedrock level: the only intelligible description that can justify the conclusion that a certain application is a correct application of a certain rule is equivalent to a simple restatement of the rule itself.

I am not denying this symmetry, but, in my opinion, McDowell has not paid enough attention to certain differences. Let us return to the case of intention. I think there is still a tension within McDowell's views on this topic. On the one hand, he has rejected, rightly in my view, any "no-fact-of-the-matter" account of intention based on Kripkean considerations. In fact, we can say that he has made this rejection the basis of his consistent attacks on deflationary

accounts of intention. On the other hand, he accepted, in his influential 1984 paper, certain parts of the Kripkean argument: basically that there is no hope of accounting for intention in terms of dispositions, simply because the notion of intention is normative [McDowell (1984)]. And this idea, the idea that the notion of intention is a normative notion, in the very same sense in which the notion of meaning is a normative notion, has systematically been assumed by McDowell since then [McDowell (1998b) and (1998c)]. Against this parallelism, I will argue that the notion of intention is not normative in the sense in which the notion of meaning is normative.

Let us consider only basic third person attributions of intention to make the discussion simpler. If Kripkean scepticism or neo-Kripkean deflationary accounts have to work for first person attributions, they have to work for basic third person attributions too. And McDowell will agree, I would guess, that any self-attribution of intention can only bear a determinate content in virtue of certain connections with the kind of public events that make third person attributions possible. Then, the kind of "no-fact-of-the-matter", Kripkean argument can be expressed in third person language: in virtue of what does someone have the intention IN with the determinate content DC? Obviously, we might point out that we can see in his behaviour that he does have IN with content DC. It would, though, be useless: the philosopher's question is a question about the non-intentional fact to which the intentional fact has to be reduced. There is no such fact.

Nevertheless, it is not altogether clear that, in certain basic cases of third person attributions, we can be easily seduced by Kripke's idea that intention, as opposed to dispositions, is a normative notion. Against this, it does seem clear that, in certain basic cases, someone's behaviour quite obviously expresses a certain intention IN, simply by the fact that this behaviour makes a certain kind of disposition D obvious. Then, had he not had the disposition D, my attribution of the intention IN would be false, and vice versa. The intention to avoid the approaching car that a pedestrian's behaviour expresses is simply a kind of disposition to avoid the car. The intention to avoid the predator that the prey expresses is simply a certain kind of disposition to avoid the predator. "Not so," a Kripkean would say, "intention is a normative notion whilst a disposition is merely a non-normative fact about someone". But it is difficult to accept that this might be the right answer. There is no obvious sense in which the fact that someone has an intention has to be a normative fact about *him*, while the fact that someone has the relevant disposition is not as normative. In fact, it is a strange use of the term "normative" *in both cases*. Of course, if the attributed content is DC, then *the person who makes the attribution* — even in the cases of self-attribution — *must* accept that only when DC obtains, has the intention been satisfied. (But the same may be said of the relevant disposition to DC). Both in the case of a "mere" disposition and in the case of intention, we identify dispositions and intentions independ-

ent of what is in fact going to happen. The glass has a disposition to break. It will break in certain circumstances C, provided that something does not happen to prevent the exercise of the disposition, or provided that it does not lose the disposition. The truth conditions of a standard attribution of disposition are not fixed by what the object will in fact do in the future. In the same sense, certain basic attributions of intention are such that the agent, given certain circumstances, will act in such and such a way, provided that he does not change his mind.

In my opinion, there are many different phenomena that can play a role in some common intuitions about the normativity of intention: a) the fact that a particular intention may be the intention of following or establishing a norm, b) the fact that certain ways of expressing intention are also ways of expressing a commitment — the subject can commit himself to not changing his mind, c) the fact that not every kind of disposition to do X is an intention to do X, d) the fact that an agent can make cognitive mistakes when acting on a determinate intention, e) the fact that it could be said that there is a failure in an action that is not conducive to the satisfaction of the intention with which an agent acts, f) the fact that intentions cannot be analyzed merely in terms of dispositional, non-intentional language. None of this justifies Kripke's thesis (or McDowell's approval) that "the relation of meaning and intention to future action is normative" [Kripke (1984), p. 37]. It does not justify this conclusion because some of the phenomena I have mentioned are not proper to every possible intention and, in any case, none of them is the kind of phenomenon that Kripke or McDowell had in mind as a reason for Kripke's thesis. For what they had in mind seems to be only the logical must, in Wittgenstein's sense, the internal connection between a thought, a belief or a statement and their truth conditions, between a desire or an intention and their conditions of satisfaction, between and order and what the order orders [McDowell (1998a), pp. 235-7, 265, 270, 300-2].

The relation of linguistic meaning to action is normative — in a sense I will describe later. However, the relation of intention *per se* to future action is not normative in this sense. Obviously, not all kinds of future action would be in accordance with a previous intention, not everything would count as a satisfaction of a particular desire, and not every possible state of affairs would render a proposition about the future true. We can, if we want, call this phenomenon "normativity". But then we cannot defend that it is obvious that "mere" dispositions are not normative in this sense. It is not obvious, at all. And to show exactly where the difference is, we need to say something that goes beyond "normativity" in this sense. We might defend, for instance, that propositional attitudes have content in a different sense to which a disposition could be said to have content, that the attribution of content to an intentional agent has a completely different explanatory function from the attribution of a disposition to a physical object. But this is not a difference

that could be explained in terms of the putative general phenomenon of "normativity", if this is understood as the mere phenomenon of the logical must, as the mere phenomenon of fixing only some possible states of affairs as those satisfying the relevant propositional attitude or counting as the proper exercise of the relevant disposition.[1] The question of the normativity of intention has to be, then, clearly distinguished from the question of the re-ducibility of intentional language. On this second issue, I have no quarrel with McDowell: the internal connection between a determinate content DC, the relevant dispositions and its conditions of satisfaction can only be described in intentional terms. I will now attempt to say something more about the form of these internal connections.

III

To begin, let us notice an *epistemic* source of philosophical perplexity, when people feel the attraction of "no-fact-of-the-matter" arguments. We have an impressive epistemic ability when we identify intentions and changes of mind before the relevant actions are produced. In virtue of this ability, we are able to anticipate animal actions. This epistemic ability has grounded the idea that we need a reductive account of intentional facts. For, it is com-monly argued, if there were no non-intentional facts to which intentional facts could be reduced, then our effective ability to anticipate and control the world by relying on our perception of intentional contents in other people ac-tions (linguistic and non-linguistic) would be a kind of miracle. This bad argu-ment is one of the sources of the prevalent reductive naturalism in contemporary philosophy of mind. I think that the premises of this bad argu-ment are also operating in other different directions: on the one hand, it is this epistemic ability that creates the illusion of a super-determination of intention or meaning — either by platonic entities or by self-interpreted internal objects. Once we notice the futility of those explanations, we are tempted by the "no-fact-of-the-matter" conclusion. We consider that the very ideas of content de-termination and meaning determination are incoherent. The only relevant facts are the contingent facts that are linked to our epistemic success. This suc-cess, impressive but contingent as it is, is then interpreted as *creating the illu-sion* of meaning-determination, or content-determination. In the practice of attribution of intentions, for instance, we start with the assumption that the at-tributed content fixes the conditions under which we should recognize that the intention has been satisfied. Once we renounce a certain mythology about how this content might be fixed, it seems that the naked practice of recognizing when an intention is satisfied cannot be described as being *determined* by the requirements imposed by that content. Simply because this practice *fixes* con-tent, it is difficult to accept that it might be grounded on content.[2] So we are

trapped with: either an incoherent account of content determination or a deflationary account of intentional content.

IV

The previous dilemma should be resisted. And the key to resisting it is the sophisticated link between the epistemology of content attributions and the metaphysics of (the attributed) content. Our epistemic success in attributing intentional content — our ability to anticipate future behaviour, the relevant stability of our attributions — is related, as a condition of possibility, to a special kind of non-contingent connection that is constitutive of the hardness of the logical must. Wittgenstein's first account of it can be dated to the early thirties, and it can be traced in the second half of the first part of *Philosophical Investigations*. It is a revolutionary account that marks the complete abandonment of any empiricist account of experience: intentional attributions are possible because we can find non-accidental connections *within* the field of experience. Of course, no internal connection is different from the logical must — no internal connection is more basic than the necessities that are made by logic and grammar. There is, nonetheless an aspect of these necessities that had been typically unnoticed by pre-Wittgensteinian philosophy of language: the internal connection between a picture and its truth conditions, between an intention and its condition of satisfaction, requires the stability of certain connections between different fragments of behaviour. It requires what can be called "expressive behaviour": the pictorial properties of certain ways of acting that can only depict certain contents by being non-accidentally related to their own future development. In *Philosophical Remarks* and *Philosophical Grammar*, in the early thirties, we find certain ideas about intentionality that are at the core of the reflections on rule following in *Philosophical Investigations*. Basically, the intentional object can only be fixed *via* expressive behaviour. "Tell me *how* you are searching, and I will tell *what* you are searching for" [Wittgenstein (1975), §27. Cf. Wittgenstein (1973), I, §§86, 91, 92]. The intentional link is an internal, not a contingent, relation. It is now fixed that I am looking for a determinate entity to be found in the future. This has to be fixed quite independently of the object's actual existence — I can look for something that does not exist. Nevertheless, the object I am looking for is depicted in my way of acting *now*.

Wittgenstein's *bête noire* was the idea that the non-representational features of certain detached, mental particulars can account for intentional content. I agree with McDowell, of course, that this is not an attack on the possibility that mental particulars might be endowed with intentional content. This would be plainly self-refuting: thoughts, perceptions and images do have content. By criticising classical conceptions of intentionality, Wittgen-

stein attacked certain accounts of the way in which particulars acquire pictorial powers. In fact, about the sense of "pictorial powers", the sense in which anything can have pictorial powers at all. If we try to ground those representational powers in non-intentional features we are forced to conclude that representation is not possible. Just as we cannot explain the representational powers of a physical object in terms of its intrinsic, non-representational determinations, in a similar way we cannot ground the representational powers of any particular. The other alternative that Wittgenstein seriously considers is a kind of eliminative account: Russell's behaviourism, the kind of theory that transforms the internal link into a merely contingent connection. According to Wittgenstein, the fact that I want to eat an apple has to be fixed now; it is not something that has to be determined, as Russell's theory requires, by my future behaviour. Traditional theories of intentionality are committed to a plainly incoherent story about how intentional content is fixed by postulating entities that are representationally dead and that still require some additional source of intentional life. On the contrary, behaviourist theories ignore the requirements of the logical must, by transforming internal connections into contingent links: it would be a contingent issue what the object of my propositional attitudes is. For Wittgenstein, the desired object is not the object that gives me psychological satisfaction: whether the apple will satisfy me or not is a contingent question, to be decided in the future. It is necessarily fixed now, though, that I desire an apple: this is the hardness of the logical must [Wittgenstein (1975), §§21-2].

In the light of previous considerations, what is the best reconstruction of Wittgenstein's argument for the necessity of expressive behaviour? It is, I would say, in a move that is parallel, only parallel though, to McDowell's insistence in his 1984 paper about the necessity of not digging below the bedrock. If I am right, it is at this point — intentionality, not meaning — where the prohibition gains strength. In section VI of the paper, I will try to show why McDowell's insistence loses some of its grip simply because it arrives a little too late.

The perverted dialectical setting that, according to McDowell, is accepted by Kripkean accounts of meaning can also be naturally introduced for the question of intentionality. The apparent necessity of choosing between content indetermination and a magical kind of determination can be an unpalatable option for the classical philosopher, who defends that non-representational features of particulars can be the bearers of representational powers. The behaviourist elimination would correspond to any neo-Kripkean deflationary account. What is the parallel move — regarding intentionality — to Wittgenstein's rejection of the hidden assumption that meaning has to be grounded in interpretation? Simply the rejection of the hidden assumptions behind any reductive account of intentionality. It is not merely the parallel move: it is the move

that justifies and gives full import to the idea that there must be a way of grasping meaning that is purely reaction, rather than interpretation.

If we want to avoid the behaviourist elimination of intentionality (and Wittgenstein always accepted that the elimination of the logical must was not an intelligible option), the diagnosis of the problems related to classical accounts has far reaching consequences regarding the possibility of a reduction of intentionality. Non-intentionally specified determinations of particulars are not the kind of entities that can ground intentional content. There is a *double* role in Wittgenstein's constant obsession to compare the putative mythological representational powers of mental particulars with the representational powers of physical objects. One is, of course, the pedagogical strategy of showing that the former are as representationally dead as the latter.[3] The other is to show certain constitutive aspects of the relation of representation, of the form of any "x represents y" statement. There is an obvious mistake in the idea that non-intentionally specified features can be the bearers, without further specification, of the properties that content requires: the non-representationally specified features of any representation can never determine certain features that must be present in intentional content. There is nothing in them that might fix the degree of generality and intensionality that is proper to content. This entails that the whole idea of providing a reductive account of intentionality is misguided. Not because dead mental particulars are the only dialectical option to straight, behaviouristic elimination, but because the diagnosis of their failures shows something very important about the impossibility of finding grounds for basic attributions of intentionality. Ultimately, there can be no ground for such basic attributions. Just as it is self-refuting to try to ground basic attributions of similarity in third-man entities. Such a ground would require justifying a basic attribution of similarity on some relation of — similarity! In the case of intentionality, such a ground would require the self-refuting move of justifying a basic attribution of intentionality on some intentionally specified relation: a given interpretation of the dead non-representational features.[4]

What follows from this argument has no parallel with McDowell's requirement of a linguistic community — and I am not accusing him, at all, of ever having suggested that it should have such a parallel. But this argument does provide justification for the idea of expressive behaviour as the proto-phenomenon of content, the idea that certain basic intentional phenomena are, to use McDowell's expression, bedrock territory. Our ability to perceive intentionality in certain basic forms of behaviour cannot have, and does not need, any rational ground. We are able to perceive directly the expressed content in certain basic forms of animal behaviour. The most important thing here is that in no way can we use the previous reconstruction of Wittgenstein's argument to justify the idea that certain forms of behaviour manage to have the *same kinds of* pictorial powers that mental, detached particulars *cannot* have.

This would be self-refuting: in the sense in which certain features of certain mental or physical particulars cannot be the bearers of representational powers, nothing can bear those kinds of pictorial powers — in particular, certain non-representational features of certain particular ways of behaving cannot be the bearers of representational powers either. Simply because we are introducing a confusing picture about what it to represent. Wittgenstein's argument is not (only) about putative candidates for one of the *relata* of the relation of representation: non-representationally specified determinations of particulars. The argument is about a notion of representation that creates the illusion that such entities might be able to represent. The non-intentionally specified, determinate features of the particular behaviour of an animal do not ground our attributions of intentionality, and they do not represent either.

The conclusion is meant to be that, *in spite of this*, it would be self-refuting to insist that something is missing, that we do not have all the ground that we need: for we need none. To insist that we need it would be self-refuting. To insist that we need it would be to express a basic confusion about how representation is possible, at all. It is commonplace to distinguish between intrinsic and derivative intentionality. Mind is the *locus* of intrinsic intentionality. Pictures and words have a kind of derivative intentionality: pictures and words manage to represent in virtue of their connections to the mind. Naturalizing, reductive programmes about intentionality assume that there must be some (non-intentional) facts in virtue of which minds manage to represent. This is the unintelligible assumption that Wittgenstein's criticism of non-intentional features as bearers of representation tries to attack. We can intelligibly ask of a conventional representation (a name, a predicate, a picture, a sentence): in virtue of what does it represent what it represents? If intentionality is possible at all, there must be certain proto-phenomena about which this question cannot be asked. This is a question that cannot be asked of expressive behaviour.

We could say that our perception of intentionality in this case is not based on interpretation; it is just a matter of our reaction. This would simply be an epistemological point: it is because we react in certain ways that we can have epistemic access to the expressed content in certain basic forms of action. Nevertheless, the most important metaphysical consequence concerns the form of the attributed content. The fact that our reaction is a condition for the possibility of our grasping basic contents is the epistemological side of the metaphysical *status* of content: there is no way of deriving intentionality from non-intentional features. Content cannot be reduced to non-intentional features. The perception of basic expressions of content cannot be grounded in the perception of certain non-intentional features. This means that content can only be perceived from a framework of epistemic reactions and abilities that is not justified by the perceived content. That our system of natural reactions provides the framework of measurement for the attributed content, is

true. However, this does not mean that the attributed content is a description of such a framework. Nor does it mean that we can make sense of other, alternative and very different, ways of determining contents. These alternatives would not be conceivable as ways of measuring intentional content. What they would measure would be something different. The human system of re-actions allows us to measure contents, and determines what it is like to measure contents. It does not determine what it is like "to measure contents for us", but to measure contents. Something completely different would not be an alternative way of perceiving contents. Intentionality cannot have a hidden essence. This is the metaphysical import of the idea that representational powers cannot be grounded on non-representational features of their bearers.

It is true that, if some kinds of beings were not able to see that the behaviour of an animal depicts a certain object as object of his desire, we could only accuse them of reacting in a different way to us. We could not accuse them of making any rational or inferential mistake. A Laplacean demon, for instance, without some of our natural reactions *vis-à-vis* the world, would not be able to see any similarity in different instances of the same intentional type. (For the moment let us ignore the question whether or not this counts against the intelligibility of such a kind of mind). This does not justify the thesis that our attributions of intentionality are about our own point of view, about the set of natural, unjustified reactions that make them possible. The rejection of this thesis is simply an aspect of the rejection of the reductionist assumptions about what kind of entities representational powers are, and the kind of relation that "x represents y" specifies. It is wrong to argue, for instance, that our being unable to justify in non-intentional terms the fact that x represents y, makes "x represents y" a dubious statement: a statement that cannot be about x, or that can only get a determinate content if made relative to our own perspective onto the world. This would be simply to assume the self-refuting point of view of a reductive account of intentionality. For if we insist that the true form of our basic attributions of intentionality is "x represents y, *for us*" then we are still left without any possible account of the expression "x represents y" itself. This would be, ultimately, to destroy the very requirements of the logical must.

The crucial proto-phenomenon of intentionality is then the everyday fact that we are able to see the intentional object in certain courses of action. This ability cannot have any external foundation. *But this requires a complete transformation of the old idea that the logical must places no restriction on how the world should, in fact, be.* Facts do not justify grammar: there is no external point of view from which the similarities that our grammar determines can be justified. Nevertheless, traditional views about the logical must, which deny the autonomy of grammar, are committed to the assumption that grammar is compatible with any possible systematic combination of contingent facts. Against this, Wittgenstein's new ideas entail that grammatical

conventions depend on the systematic stability between certain contingent facts: they are still contingent, in the sense that they are not made necessary by grammar. Nevertheless, the stability of their connections cannot be systematically violated without the collapse of grammar. The fact that someone's way of behaving expresses his intention of searching for some particular object entails that certain future fragments of behaviour are not accidentally linked to the pictorial properties of his actual expression. To identify something as searching behaviour requires us to see that different fragments of behaviour are parts of the same intentional pattern. The normal subject's recognition that his intention has been satisfied is also another part of the intentional pattern, so it can be non-accidentally linked to the previous expression of the intention. The kind of dilemma that is the basis of deflationary, "no-fact-of-the-matter", accounts of intention can then be avoided: it is possible for the recognition that an intention has been satisfied to be both (i) constitutive of the (previously) expressed intention, another part of which is expressed by the expression of the intention, and (ii) a *de facto* habitual consequence of the satisfaction of the intention. There is a non-accidental link between a previous expression of the intention and the actual recognition that it has been satisfied. In the same sense as a piece of searching behaviour can depict its own future development, it is usually *successful* in the depicting when the agent will stop searching, once he has found the object of his desires.[5] This is, of course, the clue to the possibility of a coherent account of both first person and third person attributions: for it is the non-accidental link between different fragments of expressive behaviour that makes it possible for a self-attribution to be non-accidentally connected to certain dispositions.[6]

V

All internal connections are made by grammar. No internal connection is justified by facts. It is not an accident that, when describing an intentional content, we use the very same words that we use in the description of its conditions of satisfaction. The internal connection between an intention and its conditions of satisfaction is no less dependent on grammar than the internal connection between a mere disposition and its conditions of satisfaction: that was my point when I previously argued that you cannot oppose the normativity of intentions and the putative lack of normativity in dispositions, simply in virtue of the internal connection between an intention and its conditions of satisfaction. By the same argument, the internal connection between different fragments of expressive behaviour is not justified by facts. In this sense, if we define expressive behaviour in terms of internal connections between different fragments of behaviour, it is still grammar-made: it is not an accident that our descriptions of the different fragments of internally related behaviour are

themselves internally related.[7] They are not internally related because, necessarily, they follow one another. Nevertheless, their particular internal connection requires that *the fact* that they follow one another cannot be the uncommon, exceptional case. So, the basic role that expressive behaviour plays in the genealogy of content helps us to identify certain natural facts as a condition for the possibility of grammar. The fact that certain internal relations (the internal relation between different fragments of behaviour) are connected to others (the internal link between intentional content and its conditions of satisfaction), can be used to identify certain natural facts that are conditions of possibility of both: we are animals endowed by nature with certain epistemic abilities regarding the future behaviour of our fellow animals [Wittgenstein (1958), I, §647].

This is not the place to attempt a more perspicuous description of the phenomenon of expressive behaviour. It would be very important, for instance, to distinguish it from the more general phenomenon of dispositionality and/or causal modality. There is a sense of "representation" or "depiction" in which it can be said that certain animal movements depict the intention to escape, and in which a stone that is falling towards my head does not depict its own future way of falling down — in spite of the fact that I naturally anticipate it.[8] This language of content or representation, which can be used in the case of action but cannot be used to say that a stone represents its own future falling down, is linked to the perception of agents as subjects of a host of many other contents, subjects for which the non-satisfaction of their purposes or the falsity of their beliefs is a kind of failure: beliefs for instance aim at truth, in a sense in which dispositions do not aim at anything. The differential normativity, so to speak, would not be produced by the mere fact that the content of a belief is fixed by certain truth conditions, but because of the fact that there is a special failure in a belief that is not true. I will not attempt to reflect on this issue: it is, nevertheless, the relevant issue to articulate the difference between the dispositions that are proper to intentional attributions and the mere dispositions that are instantiated by falling stones. Let us now consider the expressive character alone as a basic representational feature of certain forms of action: these forms of action manage to represent certain basic intentional content by representing their own future development. Even with this minimal analysis, we can articulate certain ingredients in this notion that are crucial to understanding Wittgenstein's conception of linguistic practice.

VI

If I am right in the previous paragraphs, we now have a base for understanding Wittgenstein's reflections on following a rule from a perspective that could seem slightly different to the one that has been articulated by McDowell.

In fact, to elaborate the difference, I will need to characterise another kind of connection that is proper of a normative practice. The normative link between a rule and its correct applications can only be understood by taking into account the complicated set of propositional attitudes that are required to follow a rule. McDowell introduced the requirement of a community for linguistic meaning as a consequence of the requirement that there must be a way of grasping meaning that should not be an interpretation. Even if I think that, ultimately, he was right, something more must be said about how we can obtain this consequence in the case of linguistic meaning. Obviously the rejection of a universal requirement of interpretation cannot result in this consequence in every case of intentional content. The fact that intentional content is not grounded in a process of self-interpretation does not mean that an animal in complete solitude cannot have (express) many intentions. Even in the case of an adult human being who possesses normal linguistic competence: not all of my intentions-in-action are equivalent to a commitment towards a communal norm.

What is the crucial difference here? In my opinion, we must look for it in the kind of sophisticated intentions that are required to follow a linguistic rule. To follow a linguistic rule is a normative practice, in the sense that it constantly requires the intention to accommodate one's own behaviour to an established norm. Why is it not possible to establish a linguistic norm and to try to adjust the behaviour to it, in complete detachment of a social context? The only intentions that we can attribute to an animal in solitude, the only intentions he can express, are much more limited: the intention to shelter, the intention to escape from a fire, and so on. These intentions are such that in no way could they count as intentions to submit his behaviour to a pre-existing norm. By acting with these intentions, the animal is not in any relevant sense following a norm: there is nothing mistaken in the mere fact that he does not satisfy his intentions, or that he changes his mind. Certainly, an animal can make mistakes: he can act in ways that are not efficient for the satisfaction of his independently expressed intention. We might describe some behaviour of an animal in solitude as involving a relevant kind of mistake: perhaps we can imagine the situation in which he hides a bone under a tree, and later tries to recover the hidden bone by digging under the wrong tree. The idea of a mistake here is connected to the idea, defended in the previous section, that there is a kind of failure in the agent who has false beliefs or unsatisfied purposes. I doubt that these cases might be considered as cases of trying to adjust the behaviour to a previous commitment. Be that as it may, it is clear that the generality and conventionality of the norms involved in linguistic meaning is such that nobody can express his intention to submit his behaviour to this kind of rule except by the act of submitting himself to other people's correction. Nothing that a new born baby could do would count as an expression of the belief that adults are mistaken when giving names to colours. To have

this belief, the baby has to point first to the relevant norm. And he cannot do this without previously pointing to it by expressing his submission to it.

So, in what sense do I believe that McDowell was right when he defended that the requirement of a community for linguistic meaning should be derived from the idea that meaning is not a case of interpretation — in the sense that he gave to "interpretation"? The answer is, in my opinion, a little more complex than McDowell suggested. First, it is true that the rejection of the mythological sense of interpretation plays a crucial role in Wittgenstein's devastating criticisms of traditional theories of the logical must, of the internal connection between a propositional attitude and its intentional object. This is enough to reject all the philosophical mythology about understanding (itself a propositional attitude) that accompanies traditional conceptions of linguistic meaning. By itself, this kind of argument does not seem to provide enough ground to show that non-communal linguistic meaning is not possible. It can only provide it if we insist on another corollary of Wittgenstein's criticism to traditional theories of intentionality: because intentional content is not fixed by interpretation of non-representational features, it can only be fixed by expressive behaviour. And the content that can be so expressed is seriously limited by the features of the actual environment. Only a preexisting norm can provide the proper environment in which an animal can express his first intentions to conform his behaviour to a norm.

In my opinion, then, the special normativity of any linguistic practices has to be accounted for in terms of the very complicated propositional attitudes that are involved in the process of following them. And a community is the only context in which the attribution of this set of propositional attitudes is possible. Actual submission to other people is a condition of possibility of language. This submission requires the ability to perceive certain contents in other people's expressions. It requires the ability to perceive the non-contingent links between different fragments of human behaviour. Once we grant this ability, there is nothing mysterious in our certainty that tomorrow we will continue to agree when naming colours. And this cannot be used to argue that the only relevant fact is epistemic. I hope I have shown why we should not consider that this epistemic ability is guilty of creating the illusion of meaning determination. It is, on the contrary, a condition of possibility of determinate meaning.

VII

We must bear in mind that the constitutive features of the notion of expressive behaviour are crucial to understanding some very general aspects of Wittgenstein's later philosophy. They are crucial to understanding, for instance, its radical opposition to certain forms of metaphysical realism. They are also

crucial to understanding what is going on in the debate between McDowell and C. Wright on the issue of an anti-realist interpretation of Wittgenstein.

First, metaphysical realism. One could say, in current jargon, that the notion of expressive behaviour does not pick out an objective feature of the world. After all, if some behaviour is expressive, it is expressive in relation to certain reactions to it. Is it not true that the behaviour of the prey is expressive merely because there are other animals (predators) that react in a certain way? What is the reason for saying that certain forms of behaviour depict their own future development instead of simply saying that *we* anticipate the future of certain forms of behaviour? Such a reason does exist: if it is true that intentionality requires expressive behaviour, it cannot be true at the same time that expressive behaviour is not a real feature of the world. Without expressive behaviour, there is no intentionality. Without intentionality, there is no meaning. Without meaning, no *possible* description would be entitled to pick out the real features of the world. So, if intentionality is not a part of the real furniture of the world, one cannot have access to a description of the world in which certain features — as opposed to intentional content — are described as being a part of this real furniture.

Furthermore, we can see why McDowell has been right in his criticism of certain anti-realist interpretations of Wittgenstein. Without expressive behaviour, there is no intentionality. Without intentionality, the notion of following a rule by acting in accordance with other people becomes quite unintelligible. The accord, for an anti-realist like Wright, has to be described in terms that do not presuppose the meaning, or the rule. It does require, nevertheless, the ability to perceive the content of other people's cognitive reactions. But the notion of expressive behaviour cannot be treated with the medicine of anti-realism. Such a treatment would require either that the relevant manifestation of understanding expressive behaviour should be accessible to someone who is unable to grasp the content of expressive behaviour itself; or that this content could be derived from some of its features (physical, behavioural) that could be grasped by someone who cannot grasp the content as we do (by seeing the intention in the behaviour). These are precisely the possibilities that are excluded by Wittgenstein's introduction of the idea of expressive behaviour as a condition of intentionality.

Departamento de Filología y Filosofía
Universidad de Girona
Plaza Ferrater Mora, 1
17071 Girona, Spain
E-mail: josepll.prades@udg.es

Notes

[1] A curious aspect of Kripke's discussion is an important difference between the general principles about the putative normativity of intention, meaning and understanding he mentions and the particular case he has in mind. I do not dispute, of course, that the dispositions of the agent are not sufficient for the project of fixing the meaning of "plus" — in Kripke's dialectical setting. Simply, the relevant kind of intentions that might do the trick would have to include quite a complex norm in their intentional content. However, the general principle that meaning, understanding and intention are normative notions does not play any relevant role for this conclusion: had Kripke discussed, for instance, the issue whether the way in which someone understands "plus" might be reduced to a certain subset of his actual dispositions, then his conclusion would have looked much more unsatisfactory. For there is a sense of "understanding" in which understanding is not a factive attitude towards an independent norm, a sense in which something can be understood in different and incompatible ways by different people or by the very same person at different times, and *the way* in which someone understands something is fixed by his actual dispositions to act in a certain way in certain conditions.

[2] This is the line of argument behind, for instance, C. Wright's deflationary account.

[3] "For the purposes of our studies it can never be essential that a symbolic phenomenon occurs in the mind and not on paper" [Wittgenstein (1973), I, §59].

[4] "'That's *him*' (this picture represents *him*) — that contains the whole problem of representation. […] Well, the image, *qua picture,* can't do more than resemble him. […] In the case of the image, too, I have to write a name under the picture to make it the image of him" [Wittgenstein (1973), I, §62].

[5] "But the essential difference between the picture conception and the conception of Russell, Ogden and Richards is that it regards recognition as seeing an internal connection, whereas in their view this is an external connection" [Wittgenstein (1975), §21].

Compare Wittgenstein (1975), §§11, 16. Here, the *effective* ability to recognize is explained by the mere existence of the logical must. The difference between §21, on the one hand, and §§11, 16, on the other, is subtle, but important: once you reject that recognition of the satisfaction is grasping a third entity — something that cannot be internally connected to the mere content — the effective ability to recognize is explained by the mere fact that what is recognized is an internal connection. Recognizing itself, in the case of propositional attitudes, becomes a fragment of the intentional pattern that fixes the content of the attitude.

[6] This does not entail that self-attributions of intention or desire are "mere" expressions, that they cannot be true or false.

[7] "What characterizes all these cases is, that the definition can be used to read off the object of the expectation from the expectant behaviour. It isn't a later experience that decides *what* we are expecting. And I may say: it is in language that expectation and its fulfilment make contact" [Wittgenstein (1973), I, §92]. See also Wittgenstein (1973), I, §§95, 103.

[8] An excellent discussion can be found in Taylor (1979).

REFERENCES

KRIPKE, S. A. (1982), *Wittgenstein on Rules and Private Language*, Oxford, Basil
       Blackwell.
MCDOWELL, J. (1984), "Wittgenstein on Following a Rule", *Synthese*, vol. 58, pp. 325-63.
— (1998a), *Mind, Value and Reality*, Harvard, Harvard University Press.
— (1998b), "Meaning and Intentionality in Wittgenstein's Later Philosophy", in
       McDowell (1998a), pp. 263-78.
— (1998c), "Intentionality and Interiority in Wittgenstein", in McDowell (1998a),
       pp. 297-321.
TAYLOR, C. (1979), "Action and Expression", in Diamond, C. and Teichman, J.
       (eds.), *Intention and Intentionality: Essays in Honour of G.E.M. Anscombe*,
       New York, Cornell University Press, pp. 73-89.
WITTGENSTEIN, L. (1958), *Philosophical Investigations,* Oxford, Basil Blackwell.
— (1973), *Philosophical Grammar*, Oxford, Basil Blackwell.
— (1975), *Philosophical Remarks*, Oxford, Basil Blackwell.