

LEGAL INFERENCEALISM AND SEMANTIC INFERENCEALISM

Jaroslav Peregrin*

Academy of Sciences & Charles University, Prague

www: jarda.peregrin.cz

*** draft *** 10. 12. 08 ***

One of the recent trends in the philosophy of language and theory of meaning is the inferentialist project launched by Robert Brandom (1994, 2000, 2008), elaborating on the approach of Wilfrid Sellars (1953, 1954, 1956, 1974). This project views language as essentially a *rule-governed activity*, enabling a speaker to make meaningful utterances analogously to how the rules of chess enable a chess player to do such things as, say, checking his opponent's king or castling (see Peregrin, 2008). Inferentialism is also a theory of the nature of human reason and consequently of the nature of man: an ambitious project.

Leaving aside the feasibility of such ambitions, it is of profound interest that many of the ideas of contemporary general semantic inferentialism have already been tabled and discussed within the philosophy of law. Some of these ideas were originally used to account for the semantics of legal discourse whereas now they are serving to account for discourse as a whole, which suggests that their current employment is perhaps a mere generalization; however, others were submitted to illustrate the fact that legal discourse is specific rather than continuous with the everyday variety. In this paper we would like to survey two such discussions; before discussing an important instance of the former kind, we will mention one of the latter kind. We will try to show how the legal theorists anticipated the discussions currently running within the philosophy of language, and that given the current constellation of the debate, philosophy of language may learn many things from the philosophy of law.

* I am grateful to Tomáš Sobek, who brought my attention to the issues of legal inferentialism discussed in this paper.

PART I: MEANING AND INFERENTIAL LINKS

I.1 Ross's "Noît-cif tribe"

Ross (1957) invites us to imagine a community, which he calls the Noît-cif tribe, that maintains that

- (1) if a person has eaten of the chief's food he is *tû-tû*; and
- (2) if a person is *tû-tû* he shall be subjected to a ceremony of purification.

Ross argues that in this case the term "*tû-tû*" is redundant, for its only role is in forging the inferential link between the fact of eating of the chief's food and the obligation of undergoing the purification, which can be established directly, without the mediation of this term.

Let us call (1) the *condition of application* of *tû-tû* and (2) its *consequence of application*. As Ross notes, there may (and in reality probably will) be more than one *condition of application* (a person may be *tû-tû* not only when he has eaten of the chief's food, but also, say, when his totem animal is killed etc.) and, likewise, more than one *consequence* (a person who is *tû-tû* not only be subjected to a ceremony of purification, but also, perhaps, not be touched until the purification etc.) We restrict ourselves here to a case with one single condition and one single consequence of application just for the sake of simplicity and perspicuity.

It is plausible that the Noît-cifian language contains, besides the term "*tû-tû*", which has no straightforward translation into English, also an equivalent of the English "a person that has eaten of the chief's food" and also one of the English "a person that shall be subjected to a ceremony of purification". Most probably these, as in English, will be compound expressions, but for simplicity's sake let us assume that they are the terms "*tû-tâ*" and "*tâ-tû*", respectively. In view of the fact that, obviously, *he is tû-tû* in (1) and (2) is a shorthand for *he is properly called (in Noît-cifian) "tû-tû"*, we can express (1) and (2) in the following way:

- (1*) if a person is properly called "*tû-tâ*", then he is properly called "*tû-tû*"; and
- (2*) if a person is properly called "*tû-tû*", then he is properly called "*tâ-tû*"; and

In this way, (the applicability of) "*tû-tû*" links (the applicability of) "*tû-tâ*" to (the applicability of) "*tâ-tû*" and from this viewpoint it could be replaced by the direct inferential link from "*tû-tâ*" to "*tâ-tû*".

Let us now consider the term "*tû-tâ*". What are the conditions of *its* application? We could say that a person is properly called "*tû-tâ*" only if he is properly called "*tû-tû*", but thus we would be moving in a circle, as we have already explained, in reverse, the applicability of "*tû-tû*" in terms of that of "*tû-tâ*". There may also be other intermediary expressions to whose applicability the

applicability of "tû-tâ" can be reduced, but sooner or later we will arrive at the point where there are none, leaving us no option but to specify their applicability by citing extralinguistic conditions. Therefore, for simplicity's sake, we will assume that this is already the case with "tû-tâ" itself:

(3) if a person has eaten of the chief's food, then he is properly called "tû-tâ".

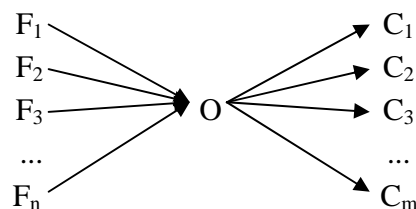
This condition is then not expressible in Noîtçifian in a nontrivial way, and the bulk of learning to use "tû-tâ" must consist in practical training. Be that as it may, let us again oversimplify and assume that this is the only *condition of application* of tû-tâ; its only *consequence of application* being

(4) if a person is properly called "tû-tâ", then he is properly called "tû-tû".

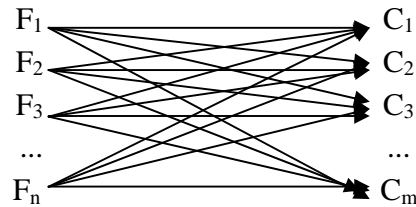
Given this, we can obviously rid ourselves of "tû-tâ" by forging a direct link from the extralinguistic situation to (the applicability of) "tû-tû". And similarly we can rid ourselves of "tû-tû" by forging a direct link from (the applicability of) "tû-tû" to the extralinguistic action. And, of course, this does not alter the fact that we may cancel "tû-tû" by forging a direct link from "tû-tâ" to "tû-tû", or, if we consider these already eliminated, from the situation to the action.

Why have we taken this detour, via "tû-tâ" and "tû-tû", if what we have finally arrived at resembles so closely our starting point? We wanted to point out that the elimination which is going on may be seen as concerning not only the special term "tû-tû", but other, more ordinary terms of Noîtçifian as well. But have we really shown it? Have we not achieved this result only by gerrymandering the usage of the ordinary terms in such a way as to make them equally artificial as "tû-tû", which is enough to provide for their elimination?

The gerrymandering admittedly happened, but how essential was it? How might the actual usage of "tû-tâ" in Noîtçifian differ from what is spelled out by (3) and (4)? First, there may be more conditions of applicability of "tû-tâ" than the one specified in (3), and there would probably be more consequences of application. But this constitutes no decisive difference between "tû-tâ" and "tû-tû". Also, the latter may have more than one condition and more consequences. As Ross draws the picture, the general situation would be



where the F's stand for the conditions and the C's for the consequences. Eliminating O then simply amounts to forging links between every F and every C:



This indicates that O leads to a great economy of reasoning: whereas without its help we need $n \times m$ links, with it we can make do with merely $n+m$.

Second, whereas the conditions and consequences for "tû-tû" were explicitly defined and are hence precise, those for "tû-tâ" do not have any such definition and are hence vague. Again, this does not seem to be a decisive difference. I do not think that anything in Ross's example hinges on the fact that (1) and (2) are explicitly formulated by the Noît-cifians. And moreover, the boundary between terms that are explicitly defined and those that are not, is not the boundary between the precise and the vague. A definition such as that of "tû-tû" above must – eventually – be couched in terms that are themselves not explicitly defined ("eat", "food", ...), which means that they too possess some amount of vagueness. On the other hand, even the terms without an explicit definition can play a useful role in language only if there is a contrast between at least things to which they are prototypically applicable and those to which they are prototypically not.

Third, the terms of ordinary language may be thought of as having a function that eludes adequate capture by any definition of the kind of (3) and (4) This is to say that although a term may have some conditions and consequences of application, its functioning in language is something over and above them. But what is it, then?

I.2 The nature of meaning

This brings us to the question of the nature of meaning. The number of answers tendered to this is vast; I can only discuss some of the most representative proposals. Nowadays many philosophers of language claim that meaning consists in truth conditions, and that the meaning of a general term, like, say, "fun" or "deer", is settled once we settle the truth conditions for sentences such as "This is fun" or "This is a deer". But note that this fits perfectly into Ross's scheme: we have a *condition of application* and empty *consequences of application*. And it follows that "tû-tû", far from being less meaningful than "fun" and "deer", is rendered *more* meaningful!

However, I do not believe that the purely truth-conditional account for meaning is satisfactory, not at least if it is taken to imply that the meaning of a term is settled once it is settled what it is applicable to. Can we make a word mean *rabbit* by making it applicable to and only to rabbits?

My hunch is that many philosophers (probably including Quine, 1960, thanks to whom referring to rabbits has become part of philosophical folklore) might consent to this. But consider a 'rabbit-detector', a mobile computer equipped with some sensors and an analytic unit which enables it to detect a rabbit if it faces one, and emit a kind of sound. Does this make the sound *mean rabbit*, i.e. express the concept of *rabbit* which is expressed by English "rabbit"? I do not think we can sensibly answer this question in the positive. Hence, the principal question is what must be added to 'rabbit-detecting' to make a 'rabbit-indicating sound' meaningful in the way words of our human languages are.

Many theoreticians of language may well wish to say that what makes the difference is the presence of a *mind* and a link between the English word "rabbit" and something in that mind, which is absent in the case of the 'rabbit-indicating sound' and the computer producing it. But this, I am convinced, would incur an irresponsible philosophical debt, which would be very difficult to redeem later. My point is that if we now explain the concept of meaningfulness, and hence of language, using the concept of mind, we would need, at some point, to explain the concept of mind independently of the concept of language and linguistic meaning. And whatever might have been the sins of the *linguistic turn* taken by so many philosophers of the twentieth century, it did document very clearly the arduousness of such a task.

Therefore I think we should look elsewhere for an explanation of what it takes for a sound to be meaningful. Especially, I think, we should try to explain it in terms of something that can be objectively detected, unlike the presence of something in one's mind. Hence optimally we should try to account for the difference between "rabbit" in the mouth of an English speaker and a 'rabbit-indicating sound' in the speaker of a rabbit-detecting computer in terms of what the respective entities *do*. While for the latter the *only* thing it does with the sound is to indicate the presence of rabbits, the role of the former is much more complicated.

The fundamental differences are of two kinds, the first hanging on the fact that "Rabbit" does not indicate rabbits in the sense that an English speaker would regularly emit "rabbit" whenever (s)he is confronted with a rabbit. As this is simply obvious, semantic theories often replace the claim that an English speaker regularly reacts to the presence of a rabbit by emitting "rabbit" by the claim that (s)he is merely so *disposed*, and that the disposition is not always realized. I do not think this helps: unless *to be disposed to emit "rabbit"* means nothing over and above *sometimes emit "rabbit" and sometimes not*, I do not see any reason why we should say that a person looking at a rabbit and not saying "rabbit" is disposed to say it. I think that what makes the genuine connection between the occurrences of rabbits and the word "rabbit" is the fact that if (s)he were to state "This is a rabbit" it would be *correct*.

It is important to realize that the difference between disposition and correctness or propriety is deeper than it might *prima facie* seem. Whereas *disposition* is a *psychological* (or perhaps a neurological) matter, *propriety* is an essentially *sociological* matter. Whereas you can be disposed to react to funny things by "fun" all by yourself, it does not make sense to say that your reaction is *correct* unless you are considered as a member of a community sharing a complicated

set of linguistic practices. (And it is crucial to realize that the reasons for this are not only those tabled by Wittgenstein, 1953, in his discussion of the impossibility of private language – the present concern is that the concepts of *disposition* and *propriety* belong to utterly disparate levels of description.

The second kind of difference between "rabbit" and a rabbit-indicating sound relates to what other things its producer does with it. Whereas in the latter case it is nothing (the indication is the end of it), in the former it is clearly *something*; but what is this something? In my view it is that "This is a rabbit" may not only be itself correct (in some situations), but can also *make other things correct*. Thus, "This is a rabbit" justifies "This is a mammal" or "This is not a mammoth". And some utterance may not only justify other utterances, but also justify actions. The claim "This is edible" justifies eating the thing in question.

It is important to see that this 'making correct' can go across individuals: *my* claim that this is edible might justify *your* eating it, hence the emerging network of normative links is an essentially *interpersonal* matter. It is not only that something that a person claims or does may vindicate something else that the same person claims or does; it may vindicate what *another* person claims or does. It follows that the level of description of our practices which we call *linguistic* and on which we are operating is a sociological one.

To sum up: the difference between "rabbit" and a rabbit-indicating sound consists in that (a) the link between the former one and rabbits is normative, whereas that of the latter one and rabbits is a causal one, and that (b) the former link, unlike the latter, is an entry point to a much larger game, in which it acts not only as a label for 'conditioning facts', but also as a means of stating various 'conditioned consequences'.

I.3 "Tû-tû" vs. "ownership" vs. "fun"

The reason for Ross himself performing his thought experiment was to indicate that some legal terms, such as the term "ownership", have the same status as "tû-tû". He claims that both these terms merely forge links between some factual conditions and some normative consequences. The point is that the usage of the term "ownership", at least in the legal context, is governed by conditions of the very same kind:

(1*) If A has lawfully purchased an object, ownership of the object is thereby created by him.

(2*) If A is the owner of an object, he has (among other things) the rights of recovery.

Hence Ross claims:

[T]he “ownership” inserted between the conditioning facts and the conditioned consequences is in reality a meaningless word without any semantic reference whatever, serving solely as a means of presentation.

The moral Ross draws from this exercise is that legal terms like "ownership" lack any reference and hence are not words meaningful in the sense in which 'standard' English words are. The background assumption seems to be that a term, to be truly meaningful, must obtain a reference; it is insufficient merely to have the conditions and consequence of its application.

However, this is precisely what current semantic inferentialism denies: its claim is that language is not a set of labels stuck to things in order to facilitate reference; rather its words become meaningful by being entangled within the network of inferential relationships. To be sure, these relationships do somehow involve a worldly dimension, but to see this dimension as a matter of reference is ill-conceived. (This is not to say that the concept of reference is itself ill-conceived, but reference should be seen as something more like a by-product of inference.)

Take the sentence "This is edible". It 'follows' from pointing at an apple or a sugar cube, and it follows from "This is an apple", "This is a sugar" or "This was eaten by my friend without any harm". It 'entails' eating it, and it entails "This is not poisonous" or "A man can eat this and it should not kill him". This way of talking about 'following' and 'entailing' is, to be sure, somewhat sloppy. To be less sloppy we may adopt the term "transition" introduced by Wilfrid Sellars. Sellars (1974) talks about three kinds of transitions:

- (1) Language Entry Transitions: The speaker responds to objects in perceptual situations, and in certain states in himself, with appropriate linguistic activity.
- (2) Intra-linguistic moves: The speaker's linguistic conceptual episodes tend to occur in patterns of valid inference (theoretical and practical), and tend not to occur in patterns which violate logical principles.
- (3) Language Exit Transitions: The speaker responds to such linguistic conceptual episodes as "I will now raise my hand" with an upward motion of the hand, etc.

Given this terminology, we can say that Ross's claim that "tû-tû" is redundant amounts to the claim that if we have a *language entry transition* followed by a *language exit transition*, there is no need to enter language at all, for we can make a direct transition from the extralinguistic states from which the entry transition starts to the extralinguistic activities into which the exit transition results.

In view of our earlier reasoning we can add that the same applies not just to an entry and exit transition pair, but to *any* pair of concatenable transitions (entry+exit, entry+intra, intra+intra, intra+exit), and hence that *any* term may be seen as eliminable in this sense. Meanwhile, we also claim that it does not follow that the meaningfulness of such words is somehow compromised -

on the contrary, the inferentialist stance implies that precisely this is the *only* way in which a word can be meaningful.

I.3 Eliminability of Language?

We have reached the conclusion that there is no difference in kind between the NoîtCIFian "tû-tû", on the one hand, and the NoîtCIFian "tû-tâ" and "tâ-tû" (or, for that matter, English "fun" or "deer") on the other. The resulting conception of meaning, then, seems to suggest that every sentence, and consequently every expression, is, just like "tû-tû", an eliminable middle term and that the only thing that is achieved by language is the establishment of a complicated set of links between situations and actions - a set the establishment of which without the help of language might be complicated in practice, but perhaps not impossible. Is the result, then, that language is in fact disposable?

Of course not. There are at least two kinds of reasons. First, utterances do not *only* function as mediators (though, the inferentialist insists, it is this that furnishes them with *meanings*), but may *also* constitute actions in their own right. Telling somebody that he is tû-tû may not just be stating that as he has eaten of the chief's food, he is to undergo the ritual purification, but it may also count, e.g., as an offense. Hence eliminating it would mean losing the wide spectrum of new actions available for us due to the emergence of language.

But what is truly crucial is the other reason: language provides for a substantial refinement of the classification of both situations and actions, which make many of the links established by means of linguistic items possible in the first place. Some kickings of a round thing through a square thing may not only count as such, but also as *scoring a goal*. In fact, it is characteristic of what we call human *actions* that they can only be identified and individuated in terms of normative vocabulary.

This was famously expressed, by Wilfrid Sellars (1956), using his metaphor of the *space of reasons* into which we may enter only with the help of fully-fledged language. A fully-fledged language is not only a collection of signals, but allows us to *reason*, or to play, as Brandom (1994) called it later, *the game of giving and asking for reasons*. What is important is that concepts, as distinguished 'building blocks' of human thought, are products of the mould provided by this very space. Hence, with a minimum of exaggeration, we can say that we start to live in our truly human world only when we are able to enter the space of reasons and thereby acquire the conceptual apparatus that lets us see the world as a world in our human sense of the word.

PART II: INSIDE OF RULES

II.1 Between Regulism and Regularism

In the preceding part we have stressed the fact that the inferential links conferring meanings on our sentences, and consequently on our words, are a *normative* matter - that what makes "ownership" or "fun" mean what it does is not that people tend to react to certain (linguistic or nonlinguistic) stimuli with "This is property" or "This is fun", nor that these sentences tend to prompt certain (linguistic or nonlinguistic) reactions, but rather that there are certain *rules* that render such inferences *correct* or *incorrect*. Thus we can picture our language games on the model of a game like chess: what makes the wooden pieces into *pawns* or *bishops* is not the fact that people tend to handle them in a certain way, but that the players of chess accept the rules of chess, and it is these rules that constitute the *role* of a *pawn* or a *bishop* that is subsequently instantiated by a wooden piece.

What is a rule, in this context? With chess in mind, we may be tempted to say that they are prescriptions recorded in a canonical book (like *The Rules of Chess*) - hence that they are *linguistic* objects. But from the viewpoint of our language games, this would be too restrictive - the point is that at least some of the rules of our language games must be merely implicit, rather than explicitly articulated in the form of linguistic tokens. As Wittgenstein (1953) pointed out, this must be the case in pain of an infinite regress: a linguistic token can play the role of a rule only if it is correctly interpreted, which can be done only if one conforms with the right rules of its interpretation. Hence the crux of the matter is that, to uphold the kind of inferentialism I am talking about, we must make sense of the notion of a rule implicit in practice.

A straightforward response might be that one follows a rule simply if one's behavior conforms to the given rule. According to this, one would be following the rule that lying is wrong if one does not lie. But this would engender a proliferation of the concept of rule. We would have stones following the rule of gravitation, flowers following the rules of blooming and our stomachs following the rules of digestion. While we sometimes do speak this way, it is clear that this is a sense of *following a rule* different from what we are talking about when considering the dynamics of human societies.

To identify the concept of *rule following* which is truly pertinent to what distinguishes us humans and our communities from the rest of the universe, we have to steer between the two ways of construing rules just mentioned - between what Brandom (1994) calls *regulism* (understanding rules as linguistic or other symbolic objects) and *regularism* (understanding them as mere regularities). Thus it seems that we are left with characterizing *rule following* in terms of some more complex features of the behavior of the rule followers.

Wittgenstein addresses this problem in *Philosophical Investigations* (1953, §54):

But we say that it is played according to such-and-such rules because an observer can read these rules off from the practice of the game--like a natural law governing the play.--But how does the observer distinguish in this case between players' mistakes and correct play?--There are characteristic signs of it in the players' behavior. Think of the behavior characteristic of correcting a slip of the tongue. It would be possible to recognize that someone was doing so even without knowing his language.

The key seems to be that a rule governed behavior is not a behavior which conforms to the rule unexceptionally, but rather such that its deviations from the rule elicit a specific kind of behavior on the part of other members of the community - a behavior that may be characterized as *correcting*, or *criticizing* or perhaps *sanctioning*.

II.2 Hart's internal viewpoint

This last observation - that a conduct is a deviation from an implicit rule if it elicits specific kinds of reactions - may tempt us to conclude that a community follows an implicit rule if there are ways of acting that lead to a reward and/or the deviation from those ways leads to a punishment. Thus, by saying *You should not eat of the chief's food* or *You should drive on the right side of the highway*, I express, according to this proposal, that if I do eat the chief's food I will be despised by my mates, or that if I do not drive on the right side, I will be fined.

This understanding of what is correct, or of what should be done, seems quite plausible, so it is not surprising that various theoreticians of law tried to explain law on its very basis. But H. L. A. Hart (1961) vehemently disagreed: the motivating force behind the behavior of a citizen under a rule of law is not (only) that they fear a punishment, for it may also simply be that they want to be law-obeying. To this end, he developed his theory of the internal viewpoint. As Hart's position w.r.t. law is precisely the position semantic inferentialism should - as far as I can see - assume towards the rules of our language games, I will discuss it in detail. My survey of Hart's viewpoint will be largely based on the congenial presentation given by Shapiro (2006).

Let me quote Hart's (*ibid.*, 55-56) elucidation of the idea of the internal aspect of rules at length:

When a habit is general in a social group, this generality is merely a fact about the observable behaviour of most of the group. In order that there should be such a habit no members of the group need in any way think of the general behaviour, or even know that the behaviour in question is general; still less need they strive to teach or intend to maintain it. It is enough that each for his part behaves in the way that others also in fact do. By contrast, if a social rule is to exist some at least must look upon the behaviour in question

as a general standard to be followed by the group as a whole. A social rule has an 'internal' aspect, in addition to the external aspect which it shares with a social habit and which consists in the regular uniform behaviour which an observer could record.

This internal aspect of rules may be simply illustrated from the rules of any game. Chess players do not merely have similar habits of moving the Queen in the same way which an external observer, who knew nothing about their attitude to the moves which they make, could record. In addition, they have a reflective critical attitude to this pattern of behaviour: they regard it as a standard for all who play the game. Each not only moves the Queen in a certain way himself but 'has views' about the propriety of all moving the Queen in that way. *These* views are manifested in the criticism of others and demands for conformity made upon others when deviation is actual or threatened and in the acknowledgement of the legitimacy of such criticism and demands when received from others. For the expression of such criticisms, demands, and acknowledgements a wide range of 'normative' language is used. 'I (You) ought not to have moved the Queen like that', 'I (You) must do that', 'That is right', 'That is wrong'.

The internal aspect of rules is often misrepresented as a mere matter of 'feelings' in contrast to externally observable physical behaviour. No doubt, where rules are generally accepted by a social group and generally supported by social criticism and pressure for conformity, individuals may often have psychological experiences analogous to those of restriction or compulsion. When they say they 'feel bound' to behave in certain ways they may indeed refer to these experiences. But such feelings are neither necessary nor sufficient for the existence of 'binding' rules. There is no contradiction in saying that people accept certain rules but experience no such feelings of compulsion. What is necessary is that there should be a critical reflective attitude to certain patterns of behaviour as a common standard, and that this should display itself in criticism (including self-criticism), demands for conformity, and in acknowledgements that such criticism and demands are justified, all of which find their characteristic expression in the normative terminology of 'ought', 'must', and 'should', 'right' and 'wrong'.

Hence Hart (*ibid.*, 86) summarizes:

When a social group has certain rules of conduct, this fact affords an opportunity for many closely related yet different kinds of assertion; for it is possible to be concerned with the rules, either merely as an observer who does not himself accept them, or as a member of the group which accepts and uses them as guides to conduct. We may call these respectively the 'external' and the 'internal points of view'.

As Shapiro (*ibid.*, 1164-6) summarizes Hart's attitude:

Social rules, therefore, have what Hart calls both an external and internal aspect. They have an external aspect, which they share with habits, in that most members of the group conform to the behavior. Social rules and habits are regularities of behavior. But they also have an internal aspect, in that these regularities are explained by the fact that members of the group possess a critical reflective attitude. Members of the group act *as a rule* because they accept that *there is a rule*. ... [S]ocial rules differ from habits in that the former are constituted by two types of social regularities, namely, behavioral and attitudinal, whereas habits are constituted merely by one type of social regularity.

Note that there is a parallel between Hart's account of rule following and Wittgenstein's account quoted above. Just like Wittgenstein, Hart puts stress on *criticizing* the wrong-doers. As Shapiro (*ibid.*, 1164) construes Hart's standpoint, what is needed for rule following to emerge is that (1) "deviations from social rules ... engender criticism from members of the group"; that (2) "these criticisms are taken to be legitimate"; and (3) "members of the group treat the existing pattern of behaviour as a common standard of behavior". Aside of the fact that it is not clear whether the third point claims anything nontrivial over and above what is already claimed by the first two, this is an instructive summary.

II.3 Describing vs. endorsing

Note the essential difference between this view and the sanctions-based account of rules (endorsed by people like John Austin, against whom Hart came to rebel¹). While in the former case prosecuting the wrong-doers is merely a symptom of the fact that the rules are in force, i.e. that people tend to endorse them, in the latter case it is the *principal motive* of their obeying them. If it were merely the sanctions that lay behind obeying the law, then we could construe "Doing this is correct" as a shorthand for "If you do not do this, you will (may) be punished"; however, we do not want to reduce normative attitudes to descriptive, non-normative claims.

It is important to realize that when I say "That is correct", I am engaging in one of two essentially different kinds of speech acts, according to whether I am assuming a standpoint external or internal w.r.t. the propriety I invoke. In the external case I *describe*, I say that a certain community accepts, as a matter of fact, a certain rule. In the internal case, however, I do *not* (or at least do not *merely*) describe - I *endorse* the rule, I declare that I accept it as binding.

Thus, saying that I can assume the internal viewpoint is saying that I can engage in a specific kind of activity, *viz.* *endorsing* a rule or a propriety, and hence that I can assume a certain kind of attitude to something, a normative attitude. Succinctly said, I assume a normative attitude when I hold something for correct or incorrect. However, someone may ask, what exactly does it *amount*

¹ See Shapiro (*ibid.*).

to, holding something for correct or incorrect, or for right or wrong? We have already excluded that it could reduce to doing what is deemed right and avoiding what is not. Does it, then, mean supporting those who aim at the former and opposing those who try to prevent it? Or does it involve making some pronouncement containing the words "right", "ought to" or "rule"?

There are no clear answers here because these questions are based on the mistaken assumption that the normative must be reducible to the descriptive. It may be helpful to view the situation in terms of the discussion on reducing 'normative facts' to non-normative ones. Historically, this was, for example, an issue between Max Weber (1907), who attempted a reductionist account of the legal order in a society, and Hans Kelsen (1945), for whom such a reduction was a dismantling of any substantial concept of law. Kelsen came to develop his concept of *Grundnorm*, a mythical basic norm that ultimately legitimizes any other norm².

From the viewpoint of our current concern, Kelsen's stipulation is of little help – to explain normative facts by stipulating a normative ur-fact is a clear case of *osbcurum per obscurius*. But without such stipulation the normative facts seem to hang in the air – they are not allowed to land among the ordinary, sublunary facts of nature (for natural science, which is the ultimate authority w.r.t. what nature consists of, does not report the existence of any such things), yet they somehow seem to be there. I think the solution to this riddle is to abandon seeing normative facts as *facts* in the literal sense of the word. That something is correct, or that it should be, is a matter of human normative attitudes, of the fact that people are able to assume the *internal viewpoint* as envisaged by Hart.

However, it is important to realize that this does not imply that being correct simply boils down to *being held for correct by the majority*. Though there may be some (simple) norms where indeed this is the case, it is far from the case in general. There is more than one reason for this. First, norms function as *institutions*, especially as they tend to become explicit in the form of linguistic articulation, in which case they may become effective relatively independently of people's beliefs. But even without being explicitly articulated, their institutional character may cause a situation in which their effectiveness is in the hands of a relatively small number of people. Second, norms may entail other norms, hence whoever takes one norm for valid is bound to hold other norms - its consequences - for valid, even without the person noticing it. This means that norms constitute a complicated edifice, though ultimately fully underpinned by normative attitudes.

On the other hand, this 'edifice of norms' is not a ghostly entity within the natural world. If I say that you *ought to* return the money I lent you, then I am not, of course, referring to some mysterious fiber connecting the two of us, and detectable by natural scientists if they only looked harder. I am invoking the fact that if you do not repay your debt, our society possesses coercive

² Cf. Turner (2007). (Note that the standpoint of Rouse, 2007, which Turner opposes in this paper, is quite close to the current one.)

mechanisms; but my speech act cannot be construed as simply referring to this fact. It is a peculiar speech act and we miss the crucial point if we assimilate it to description.

CONCLUSION

Philosophers of law investigated into the edifices of legal systems, which have come to be erected atop the jungle of human nature and human societies governed by "club law". The pillars of these edifices are *musts* and *shoulds*, and their arcs are the individual laws, supporting each other in a remarkable way. Philosophers and theoreticians of law saw these edifices as intrinsically different from the wilderness that their erection was to tame, and foreign to the inclinations and features of uncultivated human nature, including human language. But recently, inferentialist philosophers of language have realized that language belongs more to the foundations of the edifices, than to the jungle the edifices came to cultivate. They have realized that *musts* and *shoulds* are essential for the foundations of language in the first place, and that language already works implicitly on principles which subsequently become explicit and clear-cut in law and other systems of rules.

This means that, for inferentialists, it may be helpful to see language as a sort of edifice too, an edifice with an 'inside' that we can enter and inhabit. Thus, the metaphor of entering within a system of rules, which envisages that we are able to *endorse* rules, to deliberately bind ourselves with them (an ability that already Kant held to be the hallmark of humanity), is applicable to the case of language itself. But its employment goes beyond the purposes of romantic philosophers or the later Heidegger (whose metaphor of language as "the house of being" became famous): it helps us reveal many details of the workings of our language. What is explicit, albeit often hard to penetrate, in legal systems and their language, is now being discerned as implicit in the whole of language.

I am convinced that this is what ratifies the metaphor of us humans - *qua* thinking, rational, free, and responsible creatures - as creatures that are capable of building such 'insides'. Semantic inferentialism has it that human language differs from any other communicating system of any other species precisely because it is able to lead us into the *space of meaningfulness*, which totally transfigures the way we live in our world.

References

- Brandom, R. (1994): *Making It Explicit*, Harvard University Press, Cambridge (Mass.).
- Brandom, R. (2000) *Articulating Reasons*, Cambridge, MA: Harvard University Press.
- Brandom, R. (2008) *Between Saying and Doing*, Cambridge, MA: Harvard University Press.
- Hart, H. L. A. (1961): *The Concept of Law*, Clarendon, Oxford.
- Kelsen, H. (1928): *Die Philosophischen Grundlagen der Naturrechtslehre und des Rechtspositivismus*. Charlottenburg: Pan-Verlag Rolf Heise; English translation *General theory of law and the state*, Cambridge, MA: Harvard University Press, 1945.
- Peregrin, J. (2008): 'An Inferentialist Approach to Meaning', *Philosophy Compass*, to appear.
- Quine, W.V.O. (1960): *Word and Object*, MIT Press, Cambridge (Mass.).
- Rouse, J. (2007): 'Social Practices and Normativity', *Philosophy of the Social Sciences* 37, 2007, 46-56.
- Sellars, W. (1953) 'Inference and meaning', *Mind*, 62: 313-38.
- Sellars, W. (1954) 'Some reflections on language games', *Philosophy of Science*, 21: 204-28.
- Sellars, W. (1956): 'The Myth of the Given: Three Lectures on Empiricism and the Philosophy of Mind', Feigl, H. & Scriven, M. (eds.): *The Foundations of Science and the Concepts of Psychology and Psychoanalysis (Minnesota Studies in the Philosophy of Science 1)*, University of Minnesota Press, Minneapolis, 1956.
- Sellars, W. (1974): 'Meaning as Functional Classification', *Synthèse* 27, 1974, 417-437.
- Shapiro, S. J. (2006): 'What is the internal point of view?', *Fordham Law Review* 1975, 2006, 1157-1170.
- Turner, S. P. (2007): 'Explaining Normativity', *Philosophy of the Social Sciences* 37, 2007, 57-73.
- Weber, M (1907): 'R. Stammlers »Überwindung« der materialistischen Geschichtsauffassung', *Archiv für Sozialwissenschaft und Sozialpolitik* 24, 94-151; English translation *Critique of Stammler*, New York: Free Press, 1977.
- Wittgenstein, L. (1953): *Philosophische Untersuchungen*, Blackwell, Oxford; English translation *Philosophical Investigations*, Blackwell, Oxford, 1953.