

# Inferentialism and the Compositionality of Meaning

**Jaroslav Peregrin**

Academy of Sciences and Charles University, Prague, the Czech Republic

jarda@peregrin.cz

---

## Abstract

Inferentialism is a species of use-theory of meaning, which, however, identifies meanings neither with regularities of usage, nor with underlying dispositions, but rather with ‘rules’ of usage. This is, of course, underlain by the picture of language as an essentially rule governed enterprise. One of the frequent challenges to the inferentialist picture is that it cannot come up with a notion of meaning that would be compositional. In this paper I address this objection (as well as some other, related ones) and I show that it stems from a misconstrual of the inferentialist standpoint.

## Keywords

inferentialism, compositionality, semantics

## 1 Introduction

Theories of semantics of natural language in recent decades can, with a certain amount of oversimplification, be divided into two different kinds. Theories of the more traditional kind are based on the assumption that to explain semantics, we must start from explaining the meanings of individual words (for, after all, meanings are nothing else than some objects stood for by the words). It is this explanation that allows us to understand language and its workings. The other kind of approaches to language goes directly to the explanation of the workings of language, moving meanings into a more instrumental role; they are expected to come to play some - be it minor or major - role within this explanation.

The approaches of the latter kind are sometimes subsumed under the general heading of *use theories of meaning*. Most of the theories are concerned with the *regularities* of the employment of expressions by speakers; but I want to talk about a different species of a use theory, one which assumes (i) that it is not regularities that are relevant, (ii) that our linguistic practices are always

governed by rules and (iii) that it is the rules that are crucial for semantics. However, the idea that our linguistic practices are essentially *rule governed* may easily be misconstrued. Though the claim *is* that rules are essential for language and that rules are more than regularities, it does not involve the claim that speakers would have to comprehend some *explicit* set of rules. Rather the idea is that initiation into language involves practical mastering of a certain cluster of *unwritten* rules, which remain implicit to the practices as passed down from generation to generation. And what determines the contents of our words are some of such rules, especially our tendency to hold some ways of their employment as *correct* and others as *incorrect*.

What may help us illustrate the situation is the worn out comparison of language and chess (though we must keep in mind the essential difference consisting in the fact that unlike those of language, the rules of chess are all *explicit*). Just as the fact that a piece of wood is a *bishop* or a *rook* is exclusively a matter of the rules we choose to let it be governed by, the fact that some kinds of sounds have specific kinds of contents is a matter of the rules they have come to be governed by. The kind of rules that are deemed crucial from the viewpoint of the content of our words are the rules of inference: not of *logical* inference, but of inference in a much broader sense: the idea is that we know what *dog* means if we know that it is correct to infer *Maggie is not a dog* from *Maggie is a cat*, *No dog can live without oxygen* from *No mammal can live without oxygen* etc. and that we can ‘infer’ *This is a dog* from an appropriate situation.

Of course, this sketch of a semantic theory calls for a lot of elucidation: what is especially necessary is to explain our linguistic practices in such a way that it makes sense to say that they essentially involve rules - rules that are implicit in them - and that the rules that are crucial from the viewpoint of contents of expressions are the inferential ones. Fortunately, much of the work has been already done by Wilfrid Sellars (1949, 1953, 1974) and Robert Brandom (1994). We are going to rehearse the view of language envisaged in their writings in the first part of this paper. (However, I want to stress that nothing in this paper hinges on the fact that the version of inferentialism put forward here is quite faithful to the views of the founding fathers.) The most important theses inspiring the view are, let me repeat, that (i) a meaning is not an object labeled (stood for, represented ...) by an expression; and that (ii) meaning is normative in the sense that to say that an expression means thus and so is to say that it should be used so and so.

This position usually provokes two kinds of objections. First there are general objections towards the approach of meaning as a matter of rules - i.e. to the very idea of normativity of meaning. I address these objections in a separate

paper.<sup>1</sup> Besides this, there are objections targeted more specifically at inferentialism. Probably the most discussed specimen of such objections is the objection - repeatedly raised especially by Jerry Fodor and Ernest Lepore and others - to the effect that though meanings should be compositional, the compositionality of inferential roles is unattainable. This is the kind of objection I am going to deal with here.<sup>2</sup> (Hand in hand with this objection then go various allegations of circularity of inferentialism, which we will also discuss.)

## 2 Inferences vs. Inferential Rules

Let us start with a survey of the inferentialist standpoint. Inferences, understood as a kind of *mental action*, can be right or wrong in the pragmatic sense of helping or not helping us achieve what we want. However, they can be also right or wrong in a different sense: they can be also right or wrong as measured by the constitutive rules of the enterprise; for according to inferentialism language is based on a set of rules that make any meaningful talk possible in the first place. Hence the key concept of the kind of inferentialism I am going to defend here is that of *inferential rule*, not that of *inference*.

This is important to keep in mind, for it helps us avoid a common confusion. What we will talk about is a *normative* variety of inferentialism, which is different from its *causal* varieties, which does not focus on rules, but directly on the factual inferences carried out by speakers. It is the latter which some philosophers appear to have in mind when they talk about *inferential role semantics* (Boghossian 1993; Fodor & Lepore 1993); but in this paper I will reserve the word *inferentialism* for the normative version (and in fact it was coined, by Brandom, in this very sense).

To illuminate the situation, we can start exploring the parallel between language and chess. Chess can be played *rightly* in two different senses: (1) in the sense of not violating the rules of chess; and (2) in the sense of playing skillfully and tending to beat one's opponents. It is the former sense that is constitutive of the very game of chess – it is the *rules* of chess which make it possible to play chess at all (hence to play chess wrongly in the first sense means not to play it at all; and playing either rightly or wrongly in the second sense presupposes playing rightly in the first.)

---

<sup>1</sup> See Peregrin (in prep.)

<sup>2</sup> Some of the issues discussed in this paper are addressed in greater detail in Peregrin (2001).

Inferential rules, of course, concern what we should do when we communicate; but in fact it may be helpful *not* to see them as telling us what to do, for this may be potentially severely misleading. They dictate what is *allowed*; or (the other side of the same coin) what is *prohibited*. Therefore we should not see them as *commands*, but rather as *constraints*. An inferential rule does not tell us directly what to do (save in the sense of dictating what we are *allowed* to do), but rather what *not* to do. To say that  $A$  is inferable from  $A_1, \dots, A_n$  is *not* to say that whoever asserts (thinks, ...)  $A_1, \dots, A_n$ , should also assert (think, ...)  $A$ ; but rather that it is incorrect to assert  $A_1, \dots, A_n$  and to *deny* (i.e. to preclude the possibility of asserting)  $A$ .<sup>3</sup>

The rules can be seen as furnishing individual sentences, and consequently individual words, with *roles*, which then, from the inferentialist viewpoint, appear as meanings. Hence, not every inference is constitutive of meaning (in fact none is!) – only inferential *rules* are so constitutive. It is not the case that expressions have their inferential roles *in force of* having certain meanings; having the inferential role *is* having the meaning. Again, the situation has a parallel in chess: it is the rules of chess that make a piece used to play the game a *pawn*, a *bishop*, a *king* etc. It is not its makeup, but exclusively the role conferred on it by the rules according to which we decide to treat it that provides the piece with its ‘value’. It makes no sense to say that what we subject to rules are *already* pawns, bishops etc. – the pieces acquire the values *via* being subjected to the rules.

The picture of inferentialism, as drawn so far, is in a sense self-contained; but it is still ‘superficial’, in a very specific sense of the word. It describes a ‘surface’ layer of the functioning of rules. To really understand it, we need to ply an underlying level, to see that the mechanisms animating the surface level are *essentially social*<sup>4</sup> – and hence that concepts such as *rule* and *inference*, *sanction* and *reward*, etc., as understood within our kind of inferentialism, are essentially a matter of social interaction; and if we do not see them as such, we are inevitably missing the point of this kind of enterprise. (This also leads to the conclusion that meaning is to be sought seeing language as a *social institution*, rather than, say, a psychological or biological reality.) However, before we elaborate on this, let us look at some examples.

<sup>3</sup> See Peregrin (2006c) for more details.

<sup>4</sup> And it is not inappropriate to see their relationship to the underlying social mechanism as analogous to the relationship between the pressure or temperature of a gas and the underlying kinematics of its molecules.

### 3 Examples

Let us first consider an instance of an expression often considered as ‘logical’, namely the connective *and*. Within the systems of logic, this connective is usually regimented as the conjunction  $\wedge$  and is taken to be governed by the following inferential pattern:

$$\frac{A \quad B}{A \wedge B} \qquad \frac{A \wedge B}{A} \qquad \frac{A \wedge B}{B}$$

This can be read as reflecting the fact that to understand *and* it is necessary and sufficient to understand that for all sentences  $A$  and  $B$ , it is correct to assert  $A$  and  $B$  if and only if it is correct to assert  $A \wedge B$ .

Now though it is beyond doubt that the role of *and* within English is slightly more complex (in some cases, for example, it does not behave symmetrically as  $\wedge$  is bound to – viz. *He started to drink and got divorced* vs. *He got divorced and started to drink*), it seems to be obvious that this inferential pattern does capture the central aspects of the content of *and* (and if we want to account for the more marginal aspects, we can turn to some more complicated analyses such as those given within various versions of dynamic logic.<sup>5</sup>)

Hence this is the case where an inferentialistic treatment is quite uncontroversial. To be sure, things may start to appear a little bit more murky already before we abandon ‘logical’ vocabulary. Consider, for example, the English connective *if... then ...*. Most English speakers would probably agree that the sentence *if  $A$  then  $B$*  together with  $A$  in general gives us  $B$ ;<sup>6</sup> i.e. that this connective is characterized by the rule which in logic is called *modus ponens*:

$$\text{(MP)} \quad \frac{A \quad \text{if } A \text{ then } B}{B}$$

However, this may be the only extent of the agreement. The explication of the connective, material implication, within classical logic is characterized, in effect, by two additional rules:

$$(\rightarrow_1) \quad \frac{\neg A}{A \rightarrow B}$$

$$(\rightarrow_2) \quad \frac{B}{A \rightarrow B}$$

<sup>5</sup> See, e.g. Dekker (2000).

<sup>6</sup> Though even this requires some qualification: see, e.g. McGee (1985). See also Peregrin (2007).

But it is notorious that if we project these rules back on *if... then...* (mapping  $\neg$  onto English negation), we end up ascribing it an inferential behavior so deviant that  $(\rightarrow_1)$  and  $(\rightarrow_2)$  have earned the title of *paradoxes of implication*. Hence they are definitely not candidates for completing the inferential pattern characteristic of the English connective. Seeking a more adequate completion, we may learn from more sophisticated attempts to explicate conditionals within logic (not all of them, to be sure, originally motivated by the desire to get a grip on the semantics of *if... then...*). One such attempt, well known from Gentzenian natural deduction, consists in adding a generalized kind of inferential rule according to which we infer *if A then B* not from another sentence or sentences, but from what is itself an inference, namely the inference of *B* from *A*, or a justification of *B* in terms of *A*. The usual schematic articulation is

$$\frac{[A]}{B}$$

*if A then B*

Could this be an adequate completion of the pattern governing *if... then...?* This may depend on what exactly we consider to be an inference or justification (and without this clarification the proposal is essentially incomplete). Is, for example, a justification of *B* in the course of which we do not really employ *A*, a justification of *B in terms of A*? It seems that in one sense it is, in another it is not. Anyway, proposals based on the framework of natural deduction which I think do approximate the real functioning of *if... then... exist*.<sup>7</sup>

Another, parallel, attempt at an account for conditionals that we may consider using for the purposes of the explication of the inferential role of *if... then...* is offered by the axiomatic calculi of strict or relevant implication.<sup>8</sup> These restrict themselves to inferential rules in the usual sense (leading from sentences to sentences) and avoid the paradoxes of implication by characterizing the inferential behavior of implication by weaker and more sophisticated patterns than the one constituted by  $(\rightarrow_1)$  and  $(\rightarrow_2)$ . Does the pattern offered by a relevant logic (or another logical system concentrating on conditionals) provide for a wholly adequate explication of that governing the English *if... then...?* Again, I think some of them do (the plural is a matter of the fact that

<sup>7</sup> See, e.g., Tennant's (1997) relevant intuitionist implication.

<sup>8</sup> The former were introduced by C.I. Lewis (see esp. his Appendix to Lewis and Langford, 1932) and subsequently led to the establishment of modal logic; the latter are due especially to Anderson and Belnap (1975).

the semantics of *if ... then ...* is not wholly determinate), but to assess the adequacy in details is a nontrivial empirical problem.<sup>9</sup>

Now let us move away from ‘logical’ vocabulary and consider an empirical term, say *dog*. It is often assumed that what gives this word its semantics is its relation to dogs, the fact that it is a representation of ‘doghood’, that it is used to refer to dogs, etc. The inferentialist has two qualifications: first, this relationship is not at a word-object level, but at a sentence-situation level (for it is primarily a sentence, and only derivatively a word, that has an inferential role). Thus, the word *dog* may be related to dogs only via sentences such as *This is a dog* (for it is only sentences that can be used to make a move within the relevant language games).

Second, the relation is not a matter of a co-occurrence, but rather of propriety. What ties the sentence to dog-featuring situations is not the fact that speakers tend to use it in such situations (in fact these may be relatively rare), but rather that it is correct to use it so (which in turn is a matter of the speakers’ tending to take it as correct – but the level of using a sentence and taking its usage as correct are as different as moving chess pieces and holding the moves for correct).

Anyway, in the case of an empirical expression (unlike a logical one) we must understand inference broadly enough to be able to encompass, as it were, an ‘inference’ from a situation to an utterance. (More precisely: the fact that it is correct to use the sentence in such a situation. Some inferentialists would probably refuse to talk about ‘inferences’ here and would seek ways to describe the situation in terms of inferences proper;<sup>10</sup> but in the current context putting it in my way may be illuminating.) But this is not enough – what distinguishes, according to the inferentialist, *saying that this is a dog* from a mere reaction to the presence of a dog (such as a specific sound emitted by a cat on encountering a dog) consists in that the sentence, besides being normatively related to a kind of situation, must be caught within the network of inferences proper – must be inferable from other sentences and other sentences must be inferable from it.

---

<sup>9</sup> Some relevant truly empirical studies of the ways people tend to understand conditional grammatical constructions do exist – see, e.g., Verschueren, Schaeken & d’Ydewalle (2005) or Counihan (2008). However, their results are not in a form amenable easy to (dis)confirm the usability of the various logical conditionals to the means of natural language.

<sup>10</sup> Thus, Brandom would describe the special status of an ‘observation sentence’ *This is a dog* not in terms of the fact that it is, as it were, ‘inferred’ from certain situations, but rather in terms of the fact that speakers can treat its utterances by certain persons (‘reliable observers’), so to say, *veridically* – they can use reports of such utterances as a premise of inferences yielding them knowledge.

To see what makes the difference, what makes our saying that this is a dog an *application of the concept* of dog, we must look at what is constitutive of a concept. And the inferentialist's claim is that it is the inferential role of the corresponding word *vis-à-vis* other words: the fact that *being a dog* entails *being a mammal*, is incompatible with *being a cat*, etc. Hence what is needed for our sentence *This is a dog* to be an application of the concept is that inferences of the following kind be in force:

$$\frac{\textit{This is a dog}}{\textit{This is a mammal}}$$

$$\frac{\textit{This is a dog}}{\textit{This is not a cat}}$$

Do these inferences, together with the 'inference' from the presence of a dog to *This is a dog* exhaust the inferences governing the term *dog* and conferring its meaning on them? Surely not. Its meaning, the concept of dog, if you want, is enormously more intricate than to be a matter of these three simple rules. And pinning it down is again a matter of an (extremely intricate) empirical investigation.<sup>11</sup>

#### 4 The Communal Basis of Inferential Rules and Inferential Roles

Now let us return to the elaboration of the communal dimension of our inferentialist picture. The basic gears of the underlying communal machinery are the deontic statuses of *commitment* and *entitlement*. Take assertion. When I *assert* something (in contrast to, say, presenting it as a joke), I undertake a responsibility to substantiate it (i.e. *to give reasons for what I assert*) whenever it is seriously challenged (i.e. *if I am asked for the reasons*) – unless I want to be considered a dim-witted windbag, whose future utterances should not be taken seriously as assertions. This means that asserting, giving reasons and asking for them are inextricably connected activities. When I assert something,

---

<sup>11</sup> At this point some readers may be frustrated by my failure to offer any definite inferential patterns characterizing English words. However, keep in mind the complexity of natural language: even seemingly very simple words are governed by exuberant and intricately interdependent systems of rules. (After all, I have not seen any adequate and exhaustive characterization of the meaning of either *if... then...* or *dog*, as used in English, in *any* kind of theoretical framework.) And we must appreciate that the claim that meaning is a matter of an inferential pattern is not simply an empirical generalization; rather it is the result of an argument based on an analysis of what a human kind of meaning, or conceptual content, can feasibly consist in.



I also offer an entitlement – the entitlement for my audience to repeat my assertion deferring its substantiation to me.

Similarly we can characterize a denial. (We might, of course, want to say that denial is the assertion of a negation; but this, from the viewpoint of the inferentialist, would be putting the cart before the horse.) We treat denial as a *challenge*: the very challenge which may be understood as a request for substantiation. Taken thus, denial can be reduced to incompatibility: we take it that some assertions simply do not go together (and the negation of a statement can then be construed as a ‘minimal’ statement whose assertion is incompatible with it). On the other hand, some assertions *entail* other assertions: any substantiation of the former is easily convertible into a substantiation of the latter; hence the former can be given as a *reason* for the latter. This is to say that the basic move of the language game of *giving and asking for reasons* (Brandom’s term) is assertion, and the game is fuelled by the fact that some assertions are incompatible with others, whereas some can be given as reasons for others.

From the inferentialist viewpoint, what an inferential rule spells out is primarily not truth-preservation, but preservation of deontic statuses, prototypically commitment. (Brandom maintains that there exist several levels of inference induced by the preservation of various statutes, but let us leave this aside here and concentrate on commitment.) To say that *A* is inferable from *X* is to say that whoever is committed to all elements of *X*, is thereby committed to *A*. Being clear about the inferential structure of language is thus part and parcel of being able to ‘keep score’ of the commitments (and other deontic statuses) of one’s fellow speakers. The point is that an utterance may bring about, besides the change elicited directly by the overt assertion, many more changes covertly, because by committing myself to one claim I commit myself to other claims. Thus, by committing myself to *This is a dog* I commit myself to *This is not a cat* and many other claims. And inference is *nothing else* than this commitment-preservation – it relates a finite set of sentences to a sentence iff the commitment to all elements of the former brings about the commitment to the latter.

This way of approaching the speech acts of assertion and denial leads to a kind of pragmatics which is essentially normative: it characterizes the speech acts in terms of the kinds of rules that govern them and in terms of those changes of normative statuses of the participants of communication which they bring about. Participation in communication essentially involves score-keeping. Semantics, then, is, in effect *nothing else* than a theory of roles conferred on linguistic tokens by the rules, i.e. of the ways in which playing these tokens changes deontic statuses.

## 5 The Implicitness of Rules of Language

We have pointed out some of the similarities between language and a game like chess; but it is also important to point out the most basic *dissimilarity* between language and chess. The crucial difference between the way in which a game like chess is rule-governed and that in which a language is consists in the fact that whereas the rules of chess are explicit, those of language are, in general, not.<sup>12</sup> We can learn the rules of chess by taking a book and studying them; whereas we can acquire the (core) rules of language only by practical initiation.

Why is this and why is the difference important? For a rule to be explicit is for it to be expressed in such or another language; and it can act as a rule only for those who can *interpret* it. Even a signpost, as Wittgenstein (1953: §85) pointed out, must be interpreted

A rule stands there like a sign-post. – Does the sign-post leave no doubt open about the way I have to go? Does it shew which direction I am to take when I have passed it; whether along the road or the footpath or cross-country? But where is it said which way I am to follow it; whether in the direction of its finger or (e.g.) in the opposite one?

However, interpretation is precisely the kind of enterprise that requires rules. Thus, to follow an explicit rule one needs a further rule, and hence if every rule were to be explicit, an infinite regress would be looming.

Yet, what does it mean to follow an *implicit* rule? How can a rule exist otherwise than via being codified? What are rules, if they are not explicit directives? The first answer that may come to mind is that to follow an implicit rule is nothing over and above acting regularly. Can implicit rules be simply assimilated to regularities? (If so, then individual inferences *would*, after all, be relevant for meaning, though via some statistic aggregation.)

The inferentialist *credo* is that even *implicit* rules must be *more* than regularities – for the inferentialist would see the fact that a theory assimilates a person acting in accordance with a rule (be it an explicit rule of chess or an implicit rule of language) to a stone falling down in accordance with the law of gravitation as its *reduction ad absurdum*. The ability to bind oneself with rules, especially the rules of inference is a distinctively human ability, which is connected to the very essence of being human.

---

<sup>12</sup> Another grave dissimilarity consists in the fact that while the game of chess is ‘self-contained’, our language games are usually ‘opened to the world’ (see Peregrin 2001: §1.3); this becomes crucial as soon as we think about the roles of *empirical* words, as we saw above in case of the word *dog*.

Hence how do the rules of language exist? We have already rejected the view that they could be reduced to *regularities* of behavior, that *following a rule* might amount simply to *being in step with the other language users*. Just as playing chess cannot be reduced to moving pieces in the way others do, but rather involves treating what the others (as well as I) do as correct or incorrect, using a language presupposes, according to inferentialism, a kind of a *normative attitude* to the utterances of others (as well as myself). I *take* what they do for *right* or *wrong*, which manifests itself by (though is not reducible to) my ‘rewarding’ those who do the right things, and ‘sanctioning’ those who do not (the ‘reward’ or ‘sanction’ would likely be my bestowing upon them a *status* – ranging, say, from ‘reasonable speaker’, or, in the case of chess, ‘serious chess player’ down to ‘notorious babbler’, or ‘chess clown’ –, which may, if resonating with their evaluation by other people, become their ‘official’ status within the community in question). It is one thing to accept the rules, to take them to be ‘in force’ for a given speaker or player; and it is quite another thing to move within the space delimited by the rules.

While inferencing is an individual activity, the framework of inferential rules (which makes inferencing as such possible in the first place!) is inevitably a teamwork, a matter of mutual recognition, assessment and criticism. The changes on which the inferentialist concentrates are those of a certain *social reality* (that means a kind of reality exemplified by such entities as NATO, university positions etc. – a reality which does not exist apart from human attitudes, but which is objective in the sense of independence from the attitudes or will of any individual human). Commitments, entitlements etc. must be construed as belonging to the realm of this kind of reality.

## 6 Inferential Roles

The explication of meanings in terms of inferential roles which is one of the outcomes of inferentialism is probably the most frequent target of anti-inferentialist criticism. The most common kind of accusation in this respect seems to be that this approach to meaning is *circular*. Let us, therefore, summarize the inferentialist stance and then look at some versions of this objection.

The inferential structure of language is a matter of which sentences of the language are correctly inferable from which other sentences. What we may call the *inferential potential* of a statement is the place of the statement within the structure: it is a matter of which statements are inferable from it and which statements it is itself inferable from. Only statements have inferential potentials.

In contrast to this, every expression can be considered as having a kind of an *inferential role*; viz. its particular position within the space delimited – indeed constituted – by inferential rules. The inferential role of an expression – which the inferentialist sees as the explication of the intuitive concept of meaning – can usually be seen as generated by a relatively small number of inferential rules governing the usage of the word. Indeed the inferentialist thinks that *the only* way of making a string- or sound-type into a meaningful word is to let it be governed by such an inferential pattern.

But *what exactly* is an inferential role? Well, I do not think this question is answerable any more explicitly than the question *what exactly is a role in a theater play?* Roles are products of *decomposition* (see Peregrin, 2006b); and there is no unique way of such a decomposition. The only restriction is that the roles must always add up to that which they are the decomposition of. Roles as such are not really objects - they lack clear criteria of individuation.

However, inferential roles can be *explicitated* in various ways. We can explicate the inferential potential of a statement, e.g. as the following pair of sets of sequences of statements (where  $\vdash$  stands for the relation of inferability):<sup>13</sup>

$$(1) \text{ IP}(S) = \langle \{ \langle S_1, \dots, S_n \rangle \mid S_1, \dots, S_n \vdash S \}, \\ \{ \langle \langle S_1, \dots, S_{i-1} \rangle, \langle S_{i+1}, \dots, S_n \rangle, S_{n+1} \rangle \mid S_1, \dots, S_{i-1}, S, S_{i+1}, \dots, S_n \vdash S_{n+1} \} \rangle$$

or, alternatively, by a collection of inferential rules which uniquely determine this set. The inferential role of a word  $w$ ,  $\text{IR}(w)$ , is then an entity whose constitutive property is that the inferential potential of every complex sentence can be seen as the sum of the contributions of its parts, i.e. that the inferential potential of every sentence  $G(w_1, \dots, w_n)$  (where  $G$  symbolizes any kind of grammatical way of assembling a sentence from its ultimate parts - words) equals the result of some way of combination of inferential roles of  $w, \dots, w_n$ <sup>14</sup>

$$(2) \text{ IP}(G(w_1, \dots, w_n)) = G^*(\text{IR}(w_1), \dots, \text{IR}(w_n)), \text{ for every } n\text{-tuple of words } w_1, \dots, w_n \text{ and every grammatically possible way } G \text{ of putting them together into a sentence.}$$

Thus roles are given merely through an ‘implicit definition’ – and just as Quine (1969, p. 45) claims that “there is no saying absolutely what the numbers

<sup>13</sup> See Peregrin (2006a, 2006c) for elaborations of the technical aspect of this issue.

<sup>14</sup> Note that it is not excluded that two sentences sharing the same inferential potential (the same sentences are inferable from them and they are inferable from the same sentences) differ in inferential roles. The reason is that a statement is not only a move in a language game, but also a part of other, more complicated statements. In this sense a sentence has both what Dummett (1973) called the *freestanding sense* (and what we see as the inferential *potential*) and what he called an *ingredient sense* (and what we see as the inferential *role*).

are, there is only arithmetic”, we can claim that *there is no saying absolutely what inferential roles are, there are only inferences (& compositionality)*.

It seems that the only feasible way to secure (2) is via (3) and (4), where the latter is nothing else than the principle of compositionality for roles, whereas the former amounts to the requirement that roles of sentences are at least as fine-grained as their potentials.

- (3) there is a function  $F$  so that  $IP(s) = F(IR(s))$  for every sentence  $s$   
 (4) for every grammatical rule  $R$  there is a function  $R^*$  so that  $IR(R(e_1, \dots, e_n)) = R^*(IR(e_1), \dots, IR(e_n))$  for every expressions  $e_1, \dots, e_n$  from the domain of  $R$ .

The technical aspect of inferentialism can now be seen as concentrated into the problems of characterizing the roles of expressions by means of inferential patterns (on the background of the assumption that such patterns *must* obtain, for it is only via them that expressions acquire inferential roles in the first place) and finding suitable explications of the roles.<sup>15</sup> (Hence our earlier Quinean *dictum* that “there is no saying absolutely what inferential roles are” is not to be read as implying that there is no point in *explicating* inferential roles!) We should also add that in the case of *empirical vocabulary* the relevant role naturally cannot be a matter merely of inferences in the narrow sense of the word (in which inference is a statements-statement matter) – we would have to admit also some ‘inferences’ which would be a situation-statement, or statement-action matter.

However, someone might object, all of this seems to result in the conclusion that inferential roles are rather ghostly entities, with no sharp boundaries and no clear status. Would it not be better to avoid altogether this kind of spectre? The answer to this objection is that from the perspective of her view, the inferentialist would not really mind avoiding them. What she sees as the unavoidable foundation of semantics are, to repeat, *inferential rules* and perhaps inferential patterns constituted by clusters of the rules. However, all of us – simply as a matter of fact – tend to perceive semantics as a matter of values of individual words adding up to the semantic values of complex expressions, sentences and supersentential wholes; just as we tend to perceive a game of chess as a matter of conspiracy of the powers of individual pieces. Hence the inferentialist feels obliged to account for this. She, though, sees any explication of individual meanings in the form of genuine (e.g. set-theoretical) objects as an essential idealization, not only because it does away with vagueness, fuzziness and the like, but also because it captures as objects something that is not really very much object-like.

<sup>15</sup> In Carnap’s and Quine’s sense of “explication” as the replacement of a pre-formal, fuzzy and unclear notion by a formal and precise concept.

## 7 Inferential Potential vs. Inferential Significance

The inferential potential of a sentence  $S$ , we said, can be characterized in terms of a pair of sets, one of them containing the sentences from which  $S$  is inferable, the other one containing those which are inferable from  $S$  together with all possible kinds of collateral premises:

$$\begin{aligned} \text{IP}(S) &= \langle S^-, S^+ \rangle, \text{ where} \\ S^- &= \{ \langle S_1, \dots, S_n \rangle \mid S_1, \dots, S_n \vdash S \} \\ S^+ &= \{ \langle \langle S_1, \dots, S_{i-1} \rangle, \langle S_{i+1}, \dots, S_n \rangle, S_{n+1} \rangle \mid S_1, \dots, S_{i-1}, S, S_{i+1}, \dots, S_n \vdash S_{n+1} \} \end{aligned}$$

Given some rather modest assumptions about the nature of the inferability relation this can be simplified. First, assuming the reflexivity and transitivity of the inference relation, it can be shown (see Tennant, 2003) that  $S^-$  is superfluous in the sense that  $S^- \neq S'^-$  only if also  $S^+ \neq S'^+$ . Hence the inferential potential of  $S$  can be represented by  $S^+$  alone. Second, accepting the indifference of the order of premises of an inference and their free reusability, we can obviously identify  $S^+$  with a relation between finite (and perhaps, by extrapolation, also infinite) sets of statements and statements:

$$R_S = \{ \langle \{ S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_n \}, S_{n+1} \rangle \mid S_1, \dots, S_{i-1}, S, S_{i+1}, \dots, S_n \vdash S_{n+1} \};$$

or, equivalently (but in a more Tarskian vein), with a function mapping sets of statements on sets of statements:

$$F_S(M) = \{ S \mid \text{there is a } N \subseteq M \text{ so that } N \vdash S \};$$

Hence if  $\text{IP}(S)$ , explicated as  $F_S$ , maps  $\{ S_1, \dots, S_n \}$  on a set containing  $S_{n+1}$ , this can be read as claiming that if somebody believes that  $S_1, \dots, S_n$ , then her belief that  $S$  correctly entails the belief that  $S_{n+1}$ ; or that given the collateral commitments to  $S_1, \dots, S_n$ , the commitment to  $S$  brings about the commitment to  $S_{n+1}$ .

This brings to the fore the (obvious) fact that the consequences of a belief one acquires are influenced by other beliefs the persons happen to entertain. In this sense, ‘content’ is essentially, as Brandom puts it, ‘perspectival’: the significance of the belief that *the man over there left the room with blood on his hands* has a different significance for me, when I also believe that the room is an operation theatre where a doctor is trying to save human lives, than for somebody, who believes that there was a murder just committed in the room. Chess once more: Though the pieces have their ‘position-independent’ roles which reflect their ‘force’ (the role of the queen makes the queen a much more powerful piece than the pawn), the significance of pieces for a particular player in a particular position need not always reflect this: there are (rare) positions in which the knight is more useful than the queen.

Hence the *inferential significance* of a sentence within a particular context is something essentially different from its context-invariant *inferential potential*. However, having explicated the potential as the kind of function we did, their relationship turns out to be quite straightforward: the inferential significance of *S* within the context *C* is the value of the inferential potential of *S* for *C*. But this should not be read as claiming that potentials are prior to significances – a sentence has an inferential potential to the extent to which the employment of *S* becomes invariant across contexts, i.e. to which there emerge context-independent rules (which we explicate in terms of our function).

## 8 Is Inferentialism Circular?

Let us now look at the objection of circularity. Hinzen (2001), for instance, accuses Brandom of claiming, on the one hand, that pragmatics determines semantics, but at the same time on the other hand that semantics determines pragmatics (*ibid.*, 165). Similarly, Fodor & Lepore (2001) argue that if meaning is to be a creature of inferences, then we must specify of *which* inferences, where the only reasonable answer appears to be *analytic* inferences; and as *analytic* is nothing else than *in virtue of meaning*, we have the vicious circle: meaning is constituted by those inferences which hold in virtue of meaning.

There is no doubt that as inferentialism is based on the assumption that meanings are inferential roles, it is committed to the claim that meanings should be explained in terms of the rules governing what we do, i.e. in terms of (normative) pragmatics. However, if the inferentialist says that one draws an inference *because* an expression has a meaning or a content, then what she means is simply that given the expression has a certain role, i.e. is governed by certain rules, the move one so makes is a legitimate one. Hence in no way does she imply that meanings or contents would be prior to, and independent of, rules and hence of pragmatics. For an inferentialist, semantic content determines pragmatic significance only in this innocent sense. Judgments of correctness of individual inferences presuppose the content of the words only in the sense that they presuppose the inferential rules governing the words. (Again, the situation is no different in chess: When I say that I can move a chess piece thus-and-so because it is, say, a bishop, what I say is not that it must have been a bishop *before* it could be subjected to the relevant rules; rather I say that as the piece is governed by such and such rules, my move is a permissible one.)

Is there a feasible answer to Fodor and Lepore's question "which inferences are constitutive of meaning?" Well if the inferentialist wanted to be uncharitable,

then the answer could be simply: “none” – for it is only inferential *rules*, not inferences which can do the job. This is no excessive pedantry – the assumption that rules are more than regularities, and hence cannot be reduced to the factual episodes of inferencing is crucial. However, of course Fodor’s and Lepore’s objection does have a point (insofar as I understand it), for it can be taken as asking which *inferential rules* are constitutive of meaning. We can consider all kind of rules for inferring statements from other statements, surely not all of them being graspable as responsible for the meanings of participating terms. I may consider a rule for inferring *X is in France* from *X is in Paris*, which turns on the fact that Paris is in France, not on the meanings of words.<sup>16</sup>

Look at chess: The rule that I cannot move the king so that it would be immediately checked by an opponent’s piece and the ‘rule’ that I cannot move the queen in the same way are of different kinds. The latter ‘rule’ merely indicates that to move the queen in the described manner is not usually a promising way to win. However, in chess, due to the explicitness of its rules, the rules are unambiguous and there is a sharp boundary between rules of the former kind and ‘rules’ of the latter one. Therefore, the values of the pieces are clearly and distinctly delineated (and it cannot be, for example, unclear what the value of a piece is).

Anyway: it makes *no sense whatsoever* to ask whether it is the chicken of meaning or the egg of inferential rules that comes first. They are two sides of the same coin. From the inferentialist viewpoint, it is senseless to think about meaning as detached from the rules – just as it is senseless to think about a price as something independent from the enterprise of buying and selling.

One more illuminating parallel with chess: The values of chess pieces are exclusively a matter of the rules to which the pieces are subjected, and the rules are a matter of our treating some moves as right and others as wrong. Hence the value of a piece and our ‘normative attitudes’ to the way it is treated are two sides of the same coin - it makes no sense to say that something is, say, a king independently of the attitudes - to be a king *is* to enjoy this kind of attitudes.

---

<sup>16</sup> Fodor and Lepore take Brandom’s (in fact originally Sellars’) claim that the inference from *Lightning now* to *Thunder soon* is relevant for the meaning of *lightning* and *thunder* to imply that Brandom takes even inferences of the contingent kind as constitutive of meaning, which they see as absurd. It is, indeed, essential that meanings of empirical terms are co-constituted by inferences which can be seen as contingent, especially by those expressive of natural laws (Sellars 1948): though the fact that a certain kind of light in the sky tends to be followed by a certain kind of sound in the air is undoubtedly not a matter of the meaning of anything, where no such interconnection obtains, the concepts of *lightning* and *thunder* would not be appropriately applicable.



So how do we tell ‘analytic’ inferential rules from other ones? (Fodor and Lepore suppose that it is these other ones which Brandom calls *material*, but this is wrong. Material inferences comprise all the inferential rules which are a matter of more than just *logical* vocabulary, including, e.g. *X is a bachelor*  $\vdash$  *X is not married*, i.e. rules which are analytic if anything is.) Well, admittedly sometimes with difficulties: of course there is no *sharp* dividing line, as there is no clear demarcating line of meaning. This should not surprise us – it is a lesson Quine taught us long ago. (And, from the inferentialist’s viewpoint, it could be seen as a consequence of the implicitness, and consequent fuzziness, of the rules.)

However, the fact that there is no sharp dividing line between the ‘meaning-constitutive’ and the other rules related to a word, i.e. that there is no *sharp* dividing line between the dictionary and the encyclopedia, should not lead us to conclude that this distinction makes no sense whatsoever. We do, as a matter of fact, tend to see words as having something as meanings, and we do tend to divide rules we learn regarding the employment of a word into general, *sine-qua-non* rules and casual ones which we may fail to know without being thereby taken as ignorant of its meaning. We tend to see some rules as related to dictionaries rather than to encyclopedias and others the other way around. In short: we distinguish between strategic, context-independent rules which constitute meanings and tactical, context-dependent ones which do not.

This brings us to a possible misconstrual of the inferentialist standpoint. Discussing Brandom’s approach, Engel (2000) argues that a theory of normativity of meaning may be viable only as claiming that “the normative character of a content ... is not a direct property of the content itself, but a property which derives from the attitudes that one has taken towards the content” (p. 316). From the inferentialist viewpoint, this is unacceptable as a matter of principle: from this viewpoint, content does not exist apart from the normative attitudes, and therefore the attitudes cannot be *directed* at it. Our normative attitudes are directed always to what other people (and ourselves) *do*. (And it is the existence of these attitudes and the whole framework they institute which let the doings count as meaningful utterances.)

In short, there is no room for any gap between the rules (and the normative attitudes which underlie them) and meaning or content. Hence if Engel claims that “Brandom is right to say that the norms of the mental come from the normative or deontic attitudes that we take towards our own thoughts” (*ibid.*), then this cannot be regarded as true in more than a very metaphoric (and misleading) sense. There are no thoughts which would pre-exist the normative attitudes so that the attitudes could be attitudes towards them. What pre-exists are people’s overt doings, which change their character by being incorporated

into the normative framework, and we refer to this change by saying that they are endowed with content. (Similarly for Fodor and Lepore's, 2001, blatantly circular misconstrual of the inferentialist viewpoint: "there is nothing more to having a concept than knowing how to use it" should be changed to "there is nothing more to having a concept than knowing how to use *a word*".)

## 9 Are Inferential Roles Compositional?

Fodor and Lepore have repeatedly argued that inferential roles cannot be meanings for they are not compositional. However, given the nature of inferential roles as exposed above, this is a strange claim. Inferential role is nothing more than a contribution which an expression brings to the inferential potentials of the sentences in which it occurs; and it is only the principle of compositionality which enables us to talk about such contribution at all (*viz.* (4) above).

The point is that the talk makes sense only against the background of the picture of the contributions adding up to the potentials – hence of a picture informed, from the very beginning, by the principle of compositionality. (The inferentialist, however, denies that the picture can be taken as showing that the meanings of the parts of a whole are prior to that of the whole – she takes the principle of compositionality not as a description of an act of composition, but rather as a methodological principle allowing for the splitting up of the meaning of the whole to those of the parts.<sup>17</sup>) Hence from this viewpoint, the principle of compositionality is so deeply embedded within the inferentialist outlook that its validity for inferential roles is merely trivial.

Consider the rules of chess: they say that (given certain conditions are fulfilled) a king and a rook can castle. This is a move carried out by both the pieces together – not the result of coincident permissible moves of each of them. However, we take it that being a king involves being able to take its part within castling; and being a rook involves being able to take its one part – hence in this sense we *do* represent the castling as the result of contributions of the two pieces; where the only condition constitutive of the individuation of the contribution is that they add up to the accomplishment of the castling.

How do Fodor and Lepore substantiate their claim that the roles are *not* compositional? Consider the following passage (1993, p. 23):

Consider the meaning of the phrase 'brown cow'; it depends on the meanings of 'brown' and 'cow' together with its syntax, just as compositionality requires. ...

---

<sup>17</sup> See Peregrin (2001: Chapter 4) and especially Peregrin (2005).

But now, *prima facie*, the inferential role of ‘brown cow’ depends not only on the inferential role of ‘brown’ and the inferential role of ‘cow’, *but also what you happen to believe about brown cows*. So unlike meaning, inferential role is, in the general case, not compositional.

Suppose, for example, you happen to think that brown cows are dangerous; then it’s part of the inferential role of ‘brown cow’ in your dialect that it does (or can) figure in inferences like ‘*brown cow* → *dangerous*’. But, first blush anyhow, this fact about the inferential role of ‘brown cow’ does not seem to derive from corresponding facts about the inferential roles of its constituents. You can see this by contrasting the present case with, for example, the validity of inferences like ‘*brown cow* → *brown animal*’ or ‘*brown cow* → *non-green cow*’. ‘Brown cow’ entails ‘non-green cow’ because ‘brown’ entails ‘non-green’. But it does not look like either ‘brown’ or ‘cow’ entails ‘dangerous’, so, to this extent, it does not look like the inference from ‘brown cow’ to ‘dangerous’ is compositional.

What do Fodor and Lepore mean by the claim that “*brown cow* entails *non-green cow*”? The expressions in question are not sentences; hence they, by themselves, cannot be the relata of the entailment relation. Hence the point, presumably, is that some *sentences* whose important component is the former expression entail some other whose important component is the latter; namely that *X is a brown cow* always entails *X is a non-green cow*.

Anyway, it is surely not the case that if you happen to think that brown cows are dangerous, then this becomes a part of the inferential role of *brown cow*. What one happens to think is entirely irrelevant, for the role is a matter of inferential *rules*. Hence the question is, first, whether there is a *rule* letting us infer *X is dangerous* from *X is a brown cow*, and, second, whether it is the general, context-independent kind of rule which we tend to perceive as a matter of meaning. And it seems to be very improbable (though surely not *impossible!*) that even if there are speakers of some idiolect of English who did elevate the dependency of danger on the occurrence of brown cows to a rule, they would treat it as a rule of the meaning-constitutive kind.

Of course, there *may* be cases of rules for inferring *X is R* from *X is PQ* without there being a rule for inferring it either from *X is P* or *X is Q* (in English perhaps for inferring *X is not explosive* from *X is a fake bomb?*). Let us say that a predicate P *predicatively entails* a predicate Q if *X is P* always entails *X is Q*. Given this terminology, what Fodor and Lepore urge is that a composite predicate PQ may predicatively entail R without R being predicatively entailed by either P or Q. We may call the inferential role of an expression restricted to this kind of sentence its *predicative inferential role*. (Hence the predicative inferential role of P is a matter of which sentences of the shape *X is R* are entailed by *X is P* and which entail it.)

Now it may be the case that *predicative* inferential roles are not compositional. But so what? To establish the claim that inferential roles are not compositional by this train of thought we obviously further need the assumption that the inferential role of a predicate is its predicative inferential role. But what should make us accept this assumption?

I think that what could make us see the assumption as plausible would be the conviction that the meaning of a predicative expression *P* is established by means of establishing what falls under it; i.e. by means of fixing the truth values of all sentences of the shape *X is P* and of all utterances of *This is P* in all kinds of situations. However, this is not a conviction shared by the inferentialist; indeed one of the pillars of inferentialism is that a sentence cannot be meaningful *merely* in force of being appropriate in certain situations, i.e. being a conclusion of certain inferences from (linguistic or non-linguistic) premises; that it must be also itself capable of serving as a premise in inferences, where this later function need not be reducible to the former. If *X is R* is entailed by *X is PQ* without being entailed by either *X is P* or *X is Q*, then it *eo ipso* belongs to the inferential role of *P* that in combination with *Q* it predicatively entails *R*, and it belongs to that of *Q* that in combination with *P* it predicatively entails *R*.

Returning to chess, if one construed the role of *king* as determined solely by the ‘normal’ moves a king is allowed to make, then castling would come out as ‘uncompositional’ - i.e. the possibility of this common move of a king and a rook would not be predictable from the possibilities of the two individual pieces as encapsulated in their roles. However, precisely therefore we *cannot* restrict the role of *king* to the ‘normal’ moves - for the ability to castle is something which characterizes a king no less than the ability to make the ‘normal’ moves.

In fact Fodor and Lepore (*ibid.*) themselves conclude that the compositionality of inferential roles is trivial – though by way of what they appear to see as a *reductio ad absurdum* of the inferentialist approach. What they claim (*ibid.*, p. 26) is that, as they vividly put it, “analyticity, meaning (and compositionality), scrape out a living by doing one another’s wash”. Quite so; analyticity and meaning are two sides of the same coin; and so are compositionality and meaning. The inferentialist does not mean to provide a reduction of one of these concepts to the other ones. The compositionality of inferential roles is not a remarkable fact revealing us something deep about the foundations of the edifice of language, but rather a platitudinous result of the fact that roles are creatures of decomposition and so have the principle of compositionality embedded into their bones.

## 10 Compositionality vs. Finite Basis

McCullagh (2003) points out that Fodor and Lepore confuse “the idea that a theory [of meaning] be finitely stateable” (which he calls *finitistic compositionality*) with “the idea that the statements in such a finite theory concern only the syntactically atomic expressions in the language” (which he calls *atomistic compositionality*). While it is, according to the author, only the former that is a *sine qua non* of a reasonable theory of meaning, Fodor and Lepore unwarrantedly assume the latter.

This is an important distinction. It is one thing to insist that the semantics of any language must be ‘finitely-based’; and it is an entirely different thing to maintain that syntactically simpler expressions must be ‘semantically more primitive’ than the more complex ones. In many cases, extending meaning-assignment from a basic part to the whole of language is possible or even trivial (and this holds not only if the basic part consists of the syntactically simple expressions, but also when it consists of some basic *sentences*) – hence once we have *finitistic compositionality*, it need not be too difficult to restructure the meaning assignment to get *atomistic compositionality*.

To avoid misunderstanding – we can understand *atomistic compositionality* in two rather different senses: besides the ‘purely structural’ sense invoked by McCullagh (‘a theory of semantics can be based on assigning meanings exclusively to atomic expressions’)<sup>18</sup> there is also a ‘metaphysical sense’ (‘meanings of atomistic expressions are – in some metaphysical sense - prior to those of the complex ones’). And whereas atomistic compositionality in the *former* sense might be a requirement not essentially stronger than finitistic compositionality (it is probable that any semantic theory based on an assignment of meanings to a finite number of expressions can be transformed into a theory based on an assignment of meanings to the finite number of atoms), its metaphysical alternative is much stronger, and – what is more important – unwarranted. (And what Fodor and Lepore repeatedly indicate, what they are interested in is just the “metaphysical priority”.)

Hence the really crucial question for the inferentialist is: can there be a *finite* stock of inferential rules which confer meanings on all expressions of language? And as the inferentialist assumes that the meaning of every expression is fixed by an inferential pattern, i.e. a finite (and usually quite small) collection of inferential rules, she is committed to the positive answer to this question. And nothing of what Fodor and Lepore point out indicates that this commitment would be flouted.

<sup>18</sup> See Peregrin (2001: §4.4, 2005).

Note that this assumption does *not* involve the assumption that the meaning-conferring patterns of individual words are different or independent of each other – it is probable that there are patterns which involve more than one expression, the meanings of which are thus constituted in mutual interdependence. Thus, for example, the basic terms of arithmetic, those of zero, of the successor, of addition and multiplication gain their meaning in terms of the inferential pattern explicated by the axioms of Peano arithmetic.<sup>19</sup>

## 11 Composition and Substitution

In a later paper, Fodor and Lepore (2001) have launched a further, different kind of attack on the compositionality of inferential roles. What they claim is that the inferentialist is not even able to see language in terms of the usual kind of grammatical structure, for his only access to grammatical categories is via the concept of intersubstitutivity *salva congruentiae*, which, Fodor and Lepore argue, in fact cannot yield them. As the authors put it: “it’s not possible to define ‘Subsentential English wff’ in terms of ‘sentence’ and ‘substitution’”. In effect, the inferentialist cannot get from sentences, which she sees as the only independently meaningful parts of language, to their parts and hence cannot even articulate the principle of compositionality in a nontrivial way.

Brandom’s proposal is that “two subsentential expressions belong to the same syntactic or grammatical category just in case no well-formed sentence ... in which the one occurs can be turned into something that is not a sentence merely by substituting the other for it” (130). What makes Fodor and Lepore claim that grammatical categories cannot be delimited in this way? They write:

Consider, for example, the class of nouns (to which, presumably, singular terms *ipso facto* belong). You might suppose that they are a proper subset of the NPs; e.g., the ones that substitute for ‘John’ in, ‘John ate his own lunch.’ But that doesn’t work because \*‘Mary ate his own lunch,’ and ‘Mary’ is surely a noun. You could try identifying the nouns with the NPs that substitute for ‘John’ in ‘John ate his or her own lunch’ since ‘Mary’ is among those. But notice \*‘They ate his or her own lunch’. In fact, there is no way to represent the distinctions (masculine

---

<sup>19</sup> Elsewhere (see Peregrin 2006a) I have investigated the question in how we can carve out the standard meanings of logical constants of classical logic (as given by the usual truth tables) by means of inferential rules. Somewhat surprisingly, this is not possible – even in the case of the most elementary logical constants – if we understand the concept of inferential pattern most straightforwardly; but I have suggested that perhaps this way of understanding inferential patterns is not adequate and I have shown that if we modify it, the situation changes.

noun' / 'feminine noun') and ('singular noun' / 'plural noun') as a hierarchy; *all* the arrangements of these categories are possible. The moral usually drawn is that the singular nouns aren't a *subset* of the nouns; rather, they are the nouns that bear the feature '+ singular'. Patently, the taxonomies that feature assignments can generate are a superset of the hierarchical taxonomies. So, it's a serious mistake to rest a theory of grammatical categories on a notion of well-formedness that permits only taxonomies of the latter kind.

Well, it is clear that if we take Brandom's substitutional delimitation literally, then it is not *noun* or *singular term* which will come out as a category; categories will be finer-grained. Perhaps *masculine singular term*, *feminine singular term*, *masculine plural term*, *feminine plural term* etc; or perhaps something even more fine-grained. These categories can then be taken to form the basis of a semilattice and the usual grammatical categories will be some higher-level elements of the lattice (*a singular noun* is everything which is either a *singular masculine noun* or a *singular feminine noun* or ...; a *noun* is everything which is a *singular noun* or a *plural noun* or ...). Or, alternatively, we may see syntactic structure as a more abstract matter, resulting from disregarding morphological variations and allowing us to see *his* and *her* as two forms of the same syntactic units. Then the substitution of 'Mary' for 'John' in 'John ate his or her own lunch' *would* be possible *salva congruentiae*, for the appropriate adjustment of the form of 'him' would be seen as part of the substitution (just like, say, the adjustment which makes room for a longer word substituted for a shorter one). All in all, it is hard to see why the delimitation of grammatical categories of subsentential expressions in terms of intersubstitutivity *salva congruentiae* would be impossible.

Fodor and Lepore further argue that even if this were possible, it would not be possible to define a *particular category*, such as the category of *singular terms* in this way – for you would have to say that singular terms are those expressions which are intersubstitutive with ... a singular term. But this argument is odd (though I do find Brandom's own discussion of this point also a little bit confusing). After all, to specify that something is a king, we do not have to say that it has the same role as an already given king – we can cite the rules which govern it.

Different words and expressions do play various different kinds of roles within the enterprise of building well-formed English expressions (and let me stress that these roles are not *inferential* roles - they do not stem from inferential, but rather grammatical rules!<sup>20</sup>), the role of *singular terms* being simply one of

<sup>20</sup> We can see a linguistic role as a kind of an invariant of a linguistic property or a relation, which usually results in the fact that it is explicated as an equivalence class of expressions

the roles; and we can individuate the roles in terms of the particular kinds of grammatical rules to which a token playing it is subordinated. Nothing hinges on the fact that this is usually not easy and that it is probably impossible without mentioning other kinds of roles. (As Sellars, 1974, stresses, the most natural way of pointing out a role is via “illustrating” it - i.e. pointing out a word which already has it). This renders the enterprise in no way circular, but merely holistic.

## 12 The Notorious *Tonk*

Returning to the basic circularity objection, we cannot complete its discussion without mentioning the short and so much discussed paper in which Prior (1960/1) famously pointed out that if we allow for an unrestricted establishing of meanings via inferential rules, we open the door for a ‘pernicious’ operator making a language contradictory by its mere presence:

$$\begin{array}{l} S_1 \vdash S_1 \mathbf{tonk} S_2 \\ S_1 \mathbf{tonk} S_2 \vdash S_2. \end{array}$$

In a language containing this operator, any statement is inferable from any other (for  $S_1 \vdash (S_1 \mathbf{tonk} S_2) \vdash S_2$ ); especially any statement is inferentially equivalent to its own negation. Hence any language containing **tonk** is *eo ipso* contradictory; and it would seem that we should block the very possibility of introducing such an operator. The upshot is sometimes taken to be that letting inferential patterns institute meanings is pernicious.

However, this is no more reasonable than to conclude from the fact of the existence of poisonous mushrooms that it is not safe to eat any mushrooms. (The correct conclusion is that there is a point in seeking a way of telling the poisonous from the edible easily and reliably.) Prior’s example shows only that not every set of inferential rules can be seen as capable of establishing a meaning of a word – not that such establishment is impossible. But the separation of those patterns which are capable of conferring meaning worth its name

---

invariant in the required way, i.e. intersubstitutive without affecting the property or relation in question. In the case of grammatical categories it is the intersubjectivity w.r.t. the property of *grammaticality* (two expressions are thus intersubstitutive just in case the substitution of one of them for the other cannot turn the expression in which the substitution is made from grammatical to ungrammatical), in case of inferential roles it is that w.r.t. the relation of inference (two expressions are intersubstitutive in this way just in case the substitution of one of them for the other within an inference cannot turn the inference from correct to incorrect). See Peregrin (2006b).



from those which are not, is something to be considered. (With respect to logical vocabulary the separation – as shown by Belnap, 1962 – can be based on the concept of *conservativeness*. But it seems that conservativeness cannot be generally required of the inferential definitions of extra-logical words. Perhaps the separation is in general simply a matter of natural selection – it is clear that a language containing Prior’s **tonk** has little chance of being adopted and developed by any real community of speakers.)<sup>21</sup>

Suppose we enhance the rules of chess by the rule that whoever first moves a rook wins (as a result, the role of the ‘rooks’ within the ensuing game would substantially change - let us call them *winrooks*). In this way, we would clearly gain a game which nobody would see as worth playing. Any version of chess containing a winrook would be trivial. Does this show that chess itself is in some sense problematic, or that the roles of its pieces must be a matter of something more than the rules?

Prior appears to be convinced that only an expression which already has a content can meaningfully occur in an inferential rule – but he never says how this content is acquired. The inferentialist denies this: according to her, content is *instituted* by the rules, for what we call content is nothing else than governedness by a certain kind of rules. (Let us note that this *prima facie* perhaps not very intuitive notion of content started to surface in the writings of many semanticists – not only of Frege, Carnap, Sellars or the later Wittgenstein etc., but in a sense already earlier, of Kant or Hegel<sup>22</sup> – as the result of dissatisfaction with other – psychologist, Platonist etc. – approaches to content. Hence the inferentialist feels that if someone objects to his proposal on the score that it “flies in the face of our intuitions”, he should present and defend a concurrent proposal to show that something more “intuitive” is consistently achievable.)

Prior (1964: 191) went further and explicitly claimed that inferential roles cannot be meanings. But his reasons for this claim are rather obscure. He writes:

It is one thing to define ‘conjunction-forming sign’, and quite another to define ‘and’. We may say, for example, that a conjunction-forming sign is any sign which, when placed between any pair of sentences P and Q, forms a sentence which may be inferred from P and Q together, and from which we may infer P and infer Q. Or we may say that it is a sign which, when placed between any pair of sentences P and Q, forms a sentence which is true when both P and Q are true, and

---

<sup>21</sup> Or, if one wants to see language in a more Chomskyan manner not as something that speakers *adopt*, but as something that is inherent to them, speakers having *this* kind of language wired in would hardly outsmart and prevail over those having one of our kind.

<sup>22</sup> Cf. Brandom (2002).

otherwise false. Each of these tells us something that could be meant by saying that ‘and’, for instance, or ‘&’, is a conjunction forming sign. But neither of them tells us what is meant by ‘and’ or by ‘&’ itself.

The conviction of the inferentialist is precisely that being ‘and’ (more precisely: to mean what ‘and’ means in English) is nothing over and above being a ‘conjunction-forming sign’.<sup>23</sup> Prior’s main reason for rejecting this appears to be that “each of the above definitions implies that the sentence formed by placing a conjunction-forming sign between two other sentences already *has* a meaning” (*ibid.*). Hence Prior argues, in effect, that a formulation of inferential rules make sense only for an already meaningful language.

We have already admitted that *inferencing* can be done only within a ‘meaningful’ language, which, however, we argued, does not contradict the claim that the meaningfulness is a matter of inferential *rules*. The inferentialist is convinced that inferential rules do not presuppose, but *confer* meaning. Hence if Prior tries to ridicule inferentialists by ascribing them the claim “being able to do the tricks just *is* knowing the meaning”, then the inferentialist takes this as a *fair* enough characterization: if “the tricks” are not only moves within the space constituted by the inferential rules, but also the recognitions of moves as right and wrong, then yes, being able to do the tricks *is* knowing the meaning.

To turn to chess once more (and for the last time), it would make no sense to say: “What you can *check* is obviously a *king* not a mere piece of wood – hence you cannot formulate rules of chess unless you have pieces which already *are* kings, pawns, bishops ...”. The concepts *check* and *king* are established in mutual interdependence. In general, the inferentialist is convinced that it is only an appropriate set of rules which can open up the space in which we can play our language games, i.e. communicate and even think in our distinctively human way.

### 13 Conclusion

I conclude that there is a kind of inferentialism that can withstand the objections discussed in this paper. The objection that meaning cannot be an inferential role, for whereas the former is, by its nature, compositional, the latter is not, is based, we saw, on a misconstrual of the concept of inferential role. Inferential roles are compositional just trivially; for it is the principle of compositionality that plays an essential role in their individuation.

---

<sup>23</sup> This is obviously an oversimplification in that the function of “and” in English is actually more multiverse – for example, it is often used to express temporal succession.

Does this not render the project of inferentialism trivial? Was not the task of semantics to analyze how meanings of words add up to meanings of sentences and how the meanings of sentences determine what we do with the sentences, i.e. which kinds of inferences we draw? Inferentialism does not start from meanings, but rather from inferential rules (thus moving meanings to a position of a useful, but in principle dispensable intermediary), because it sees the rules as the ultimate ‘linguistic reality’ - as what is constitutive of our language games and hence of language. It studies how basic rules of language bear other rules and how they all determine what is correct and what is not in the realm of language.

This also preempts the objection of circularity. Inferentialism *does* aim at explaining meanings by means of inferential rules, but it *does not* aim at explaining inferential rules by means of meanings. There is a sense in which we can say that it aims at explaining *inferences* in terms of meanings, but this is the sense that it explains *correctness* of inferences in terms of inferential *rules*. This is an enterprise that is not circular in any problematic sense.

## Acknowledgement

Work on this paper was supported by the research grant No. 401/07/0904 of the Czech Science Foundation. I am grateful to Ken Turner for helpful criticism.

## References

- Anderson, Allan R. and Nuel D. Belnap. 1975. *Entailment: The Logic of Relevance and Necessity*. Princeton: Princeton University Press.
- Belnap, Nuel D. 1962. Tonk, plonk and plink. *Analysis* 22: 130-134.
- Brandom, Robert. 1994. *Making It Explicit*. Cambridge: Harvard University Press.
- Brandom, Robert. 2002. *Tales of the Mighty Dead*. Cambridge: Harvard University Press.
- Counihan, Marian. 2008. If  $p$  then  $q$  . . . and all that: Logical elements in reasoning and discourse. *Journal of Logic, Language and Information* 17: 391-415.
- Dekker, Paul. 2000. The semantics of dynamic conjunction. In J. Seligman and P. Blackburn (ed.), *Logic, Language and Computation*, vol 3:105-125. Stanford: CSLI.
- Dummett, Michael. 1973. *Frege, the Philosophy of Language*. Duckworth: London.
- Engel, Pascal. 2000. Wherein lies the normative dimension in meaning and mental content? *Philosophical Studies* 100: 305-321.
- Fodor, Jerry A. and Ernest Lepore. 1993. Why meaning (probably) isn't conceptual role. In E. Villaneuva (ed.), *Science and Knowledge*, 15-35. Atascadero: Ridgeview.
- Fodor, Jerry A. and Ernest Lepore. 2001. Brandom's burdens. *Philosophy and Phenomenological Research* 63: 465-482.

- Hinzen, Wolfram. 2001. The pragmatics of inferential content. *Synthese* 128: 158-171.
- Lewis, Clarence I. and Cooper H. Langford. 1932. *Symbolic Logic*. New York: Century.
- McCullagh, Mark. 2003. Do inferential roles compose? *Dialectica* 57: 431-438.
- McGee, Vann. 1985. A counterexample to modus ponens. [\*Journal of Philosophy\* 82: 462-471.](#)
- Peregrin, Jaroslav. 2001. *Meaning and Structure*. Aldershot: Ashgate.
- Peregrin, Jaroslav. 2005. Is compositionality an empirical matter? In M. Werning, E. Machery and G. Schurz (eds.), *The Compositionality of Concepts and Meanings*, 135-150. Frankfurt: Ontos.
- Peregrin, Jaroslav. 2006a. Meaning as an Inferential Role. *Erkenntnis* 64: 1-36.
- Peregrin, Jaroslav. 2006b. Intersubstitutivity. In M. Bilková and O. Tomala (eds.), *The Logica Yearbook 2005*, 149-163. Prague: Filosofia.
- Peregrin, Jaroslav. 2006c. Developing Sellars' semantic legacy: meaning as a role. In M. Lance and P. Wolf (ed.), *The Self-Correcting Enterprise: Essays on Wilfrid Sellars*, 257-274. Amsterdam: Rodopi.
- Peregrin, Jaroslav. 2007. Logical rules and the a priori: Good and bad questions. In J.-Y. Béziau and A. Costa-Leite (ed.), *Perspectives on Universal Logic*, 111-122. Monza: Polimetrica.
- Peregrin, Jaroslav. Inferentialism and the normativity of meaning. (*in preparation*)
- Prior, Arthur N. 1960/61. Runabout inference ticket. *Analysis* 21: 38-39.
- Prior, Arthur N. 1964. Conjunction and contonktion revisited. *Analysis* 24: 191-195.
- Quine, Willard V.O. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- Sellars, Wilfrid. 1948. Concepts as involving laws and inconceivable without them. *Journal of Philosophy* 15: 287-315.
- Sellars, Wilfrid. 1949. Language, rules and behavior. In S. Hook (ed.), *John Dewey: Philosopher of Science and Freedom*, 289-315. New York: Dial Press.
- Sellars, Wilfrid. 1953. Inference and meaning. *Mind* 62: 313-338.
- Sellars, Wilfrid. 1974. Meaning as functional classification. *Synthese* 27: 417-437.
- Tennant, Neil. 1997. *The Taming of the True*. Oxford: Oxford University Press.
- Tennant, Neil. 2003. Frege's content-principle and relevant deducibility. [\*Journal of Philosophical Logic\* 32: 245-258.](#)
- Verschuere, Niki, Walter Schaecken and Gery d'Ydewalle. 2005. A dual process specification of causal conditional reasoning. [\*Thinking and Reasoning\* 11: 239-278.](#)
- Wittgenstein, Ludwig. 1953. *Philosophische Untersuchungen*. Oxford: Blackwell. English translation *Philosophical Investigations*. Oxford: Blackwell.