

Functional Role Semantics and Reflective Equilibrium

Simone Gozzano

Università di L'Aquila

simone.gozzano@fastwebnet.it

Abstract

In this paper it is argued that functional role semantics can be saved from criticisms raised by Putnam and Fodor and Lepore by indicating which beliefs and inferences are more *constitutive* in determining mental content. The Scylla is not to use vague expressions; the Carybdis is not to endorse the analytic/synthetic distinction. The core idea is to use *reflective equilibrium* as a strategy to pinpoint which are the beliefs and the inferences that constitute the content of a mental state. The beliefs and the inferences that are constitutive are those that are in reflective equilibrium in the process of attributing mental states to others.

Keywords

Functional Role Semantics and Reflective Equilibrium

Simone Gozzano
Università di L'Aquila
simone.gozzano@fastwebnet.it

There is a large family of semantic theories that are crucially based on the notion of *role*. Roles, either of conceptual or of inferential nature, are the structural features that allow for the individuation of meanings or contents. These theories, however, have been criticised for their commitment to holism, which, according to the critics, would condemn them to some serious and perhaps fatal mistake¹. The aim of this paper is to suggest a line of defence of Functional Role Semantics (henceforth FRS) from criticisms raised by Putnam (1988) and Fodor and Lepore (1992; 2002). The core idea is to use *reflective equilibrium* as a strategy to indicate which are the beliefs and the inferences that constitute the content of a mental state. As many semantic theories centred on the notion of role, FRS can be interpreted both as a theory of meaning and as a theory of mental content. In what follow the mental version of this theory will be considered, but a linguistic extension is also possible.

Individuating FRS' conceptual core

FRS claims that the content of a mental state (or the meaning of an expression) is individuated by the role of that mental state (expression) in the overall system of an individual's thoughts (language). Along this line we may collect, at least, two different definitions. According to one: "the meaning of a mental representation is its role in the cognitive life of the agent" (Block 1999, p. 331); according to another: "the contents of thoughts are determined by their construction out of concepts; and the content of concepts are determined by their 'functional role' in a person's psychology" (Harman 1987, p. 55).²

If we consider these definitions in their details we may appreciate that they are quite different. If we interpret Block's notion of "mental representation" as indicating thoughts such as *the cat is on the mat* then the first definition individuates functional roles at the level of thoughts (and sentences). Harman's definition, on the contrary, takes functional roles as applying directly to the constituents of thoughts, that is concepts (words in the case of sentences). So, there is a difference of conceptual grain. A third definition shows a more

¹ See the recent exchange by Cohen (1999) and Heal (1999) prompted by a paper by Heal (1994).

² Recently, Greenber and Harman (forthcoming) have defined Conceptual role semantics in this way: "CRS is the view that the meanings of expressions of a language (or other symbol system) or the contents of mental states are determined or explained by the role of the expressions or mental states in thinking". It should be noticed that they have added an epistemological twist to this rather semantic thesis.

complex view of functional role, one in which both levels, the whole sentence and its constituents or, for the mental variety, the thought and its constituents, are considered. Here it is: “A crucial component of a sentence’s conceptual role is a matter of how it participates in inductive and deductive inferences. A word’s conceptual role is a matter of its contribution to the role of sentences” (Block 1986, p. 628).

So, collecting all these elements together, we have that the functional role of a complex entity (thought or sentence) is given in its participating in inferences of one kind or another, and the inferences in which it participates are determined by the role of the entities that constitute it (concepts or words). Accordingly, we can distinguish between *conceptual* role semantics and *inferential* role semantics: while the former is used to individuate the content/meaning of simple entities (concepts and lexical items), the latter can be taken to elucidate the content of complex entities (thoughts and sentences).³ In what follow, I will consider FRS as a way for encompassing both conceptual and inferential role semantics. Consequently, according to FRS, the content of a mental item (where a “mental item” can be both a concept or a thought) is determined by its *functional role* in the overall pattern of interactions with other mental items. In turn, the functional role of a mental item is expressed in terms of inferential role, in case the item is a thought, or in terms of conceptual role, in case the item is a concept. We have, then, that the content of a thought such as *the cat is on the mat* is given by the inferences, deductive, inductive or abductive, in which it occurs as premise, inferential step or conclusion. The inferences in which it occurs are determined by the concepts’ roles that figure in the thought, namely “cat”, “table” “being on x, y”. In this way, FRS is a compositional theory, so facing, as we will see, an objection raised against it.

Troubles with FRS

Now, this theory has been attacked from two different but, somewhat, complementary perspectives. According to one criticism, raised by Putnam (1988; 1990), FRS is too weak as a meaning theory; according to the other, pointed out by Fodor (1987) and Fodor and Lepore (1992; 2002) it is too strong. So, what’s wrong with this theory?

Hilary Putnam thinks that conceptual role semantics (he refers to this version of the theory but his point applies equally well to FRS) is at least an obscure thesis for the following reason. Consider the case of Ancient Greeks saying, “this is water” when looking at one river. Given

³ A somewhat different perspective is taken by using the notion of *translation* as the key feature for clarifying the whole issue. According to this view, originally held by Quine and Davidson, the meaning of an expression is taken to be individuated by its possible paraphrases or reformulations. Cf. Field (2001).

that Ancient Greeks held many different beliefs from us, and given that we want to maintain that what we share with them is reference, we should conclude that “it is not overall conceptual role that [determines meaning]” (Putnam 1988, p. 52) or, as I want to stress here, content. Conceptual role semantics is then an insufficient thesis for determining content. Its main problem is that

If one cannot even informally indicate - without using such an expression as ‘regarded by speakers as part of the meaning’ or ‘central to the meaning’ - how one could decide *which* inferences and *which* beliefs fix the meaning of a word [or, alternatively, the content of a mental state], in the sense required by conceptual role semantics, then the claims made on behalf of conceptual role semantics have virtually no content (Ibid., p. 53).

This diagnosis is due to Putnam holding that reference is a completely external factor, not reducible to any individualistic belief. The point, then, is that a role theory of meaning is *too weak* in determining meaning, failing thus its main aim.

Fodor (1987) and Fodor and Lepore, in a number of publications (1991; 1992; 1993; 2002), raise the somewhat complementary problem. In the specific version exposed by Fodor and Lepore (1992) the problem of FRS lies in its adherence to an extreme and intolerable form of meaning holism.

The difficulty can be highlighted by appealing to inferences. FRS assumes that the content of thoughts is determined by their role in inferences; and thoughts’ inferential roles are determined by the conceptual roles of their constituents. So, *at least some* roles individuate the content of thoughts. Now, according to Fodor and Lepore there is no way to indicate *which* roles individuate the content of thoughts. This entails that there is no way to tell *which* inferences determine the inferential roles of a given thought. Therefore, there is no way to say which roles individuate the contents of thoughts. But roles do play an essential role, being the individuating condition for content. So, either none of them individuate any thought or they all do. Since *at least one* of them does, and there is no way of telling which, then *all* of them do. Given this slippery slope argument FRS is caught in a radical version of holism. The only way out of this version of holism would be to make appeal to the analytic/synthetic distinction in order to distinguish *which* inferences constitute the content of a thought. But this move is apparently unavailable to FRS theorists, and supposedly to all of us, because of the acceptance of Quine’s criticism to the analytic/synthetic distinction. Therefore FRS is

doomed to holism in its extreme version, and as such this theory is *too strong* in determining meaning.⁴

The specific doubts pointed out by these three authors are quite problematic because show that FRS converges toward a very radical form of holism. It is holism, as such, that worries these authors, because of its intolerable implications. Michael Dummett has depicted a comprehensive picture of such worries.

A thoroughgoing holism, while it may provide an abstractly intelligible model of language, fails to give a credible account either of how we use language as an instrument of communication, or of how we acquire a mastery of language ... The situation is essentially similar to that of a language all of whose sentences consist of single words, i.e. have no internal semantic structure ... it becomes unintelligible how the speakers of the language could ever have come to associate ... senses with their unitary sentences, let alone to achieve the same association among different individual speakers; or how any one individual could discover the sense attached by another to a sentence, or decide whether it was or was not the same as that which he attached to it. (Dummett 1973, pp. 599-600)⁵.

Specifically, the point is that if a thought is determined by its inferential role in the overall pattern of thoughts of an individual, the possibility of two individuals having *just one* different inferential role would bring us to the conclusion that they do not share *any* thought at all, given that the contents of thoughts are individuated by *all* their inferential roles. This amounts to say that either two individuals are *identical* with respect to *all* the inferences they draw, or are disposed to draw, or they cannot be said to share any thought at all. If they cannot share any thought, then, they cannot disagree nor agree or anything. Analogously, if two different time slices of a single person are not epistemically identical, they cannot be epistemically compared, and hence the notion of change of mind disappears, as any notion concerning the acquisition and the increment of knowledge. That is why the extreme version of holism is so bad.

At the same time, holism seems to be particularly at odd with respect to the principle of compositionality, according to which the meaning of a complex expression is determined by the meaning of its components plus the way these are combined together. This principle, originally held by Frege, seems to generate a conflict with FRS in its generality, when this theory is intended as a way of individuating meanings or contents. For instance, Fodor and Lepore (2002)

⁴ I leave aside further consequences, and the ensuing responses, that Fodor and Lepore draw from their main argument, such as the nonexistence of intentional laws and the impossibility to learn a language. It must be noticed that Fodor and Lepore's critique has been applied also to so-called "space-state semantics", that is, the connectionist version of FRS. For a reply, see Churchland (1998).

argue that compositionality is a non negotiable feature of any theory of meaning, but there is no way of reconciling FRS with compositionality because according to FRS the meaning of complex expressions is not determined by the meaning of its parts. So either compositionality or FRS must go. Being compositionality non negotiable, the conclusion against FRS follows. So, either FRS is committed to the bad consequences of holism or it is unable to take into account compositionality.⁶ However, as we saw, it is perfectly possible to take FRS, in its generality, as a compositional theory: the role of complex expressions is obtained by negotiating the roles of its components so to consider just few of them as relevant in the composition of the expression, and not always the same, according to the contexts. The relevancy of the roles of the composing parts raises the same questions raised by FRS' critics: how can we select which are the appropriate roles for any given situation? So, ultimately, the problem seems to stay, and a solution to it is called for.

Is there a way out?

Both the arguments by Putnam and Fodor and Lepore make reference to a hypothetical way out from all these troubles. On the one hand, Putnam refers to the need for conceptual role semantics (but the same applies to FRS) to illustrate *which* are the inferences that are *constitutive* of the content of a certain thought without making such an individuation obscure or, worse, question begging. On the other hand, Fodor and Lepore think that the only viable principled distinction to select *which* are the constitutive roles is the analytic/synthetic one, but no FRS theorist is willing to make appeal to that distinction.

In a sense, both criticisms point to the same way. FRS can be saved if it were possible to indicate which conceptual and inferential roles are *constitutive* in determining mental content (or meaning) both with respect to complex expressions as such and with respect to the composing parts. The Scylla is not to use vague expressions as those indicated by Putnam; the Carybdis is not to endorse the analytic/synthetic distinction, as pointed out by Fodor and Lepore. I will try to wriggle through between these two dangers.

By saying which are the inferential or conceptual roles that are constitutive in determining the content of a thought we are trying to solve *the* crucial question of FRS. As Block puts it, "A crucial question for conceptual role semantics (*the* crucial question) is what counts as identity and difference of conceptual role" (Block 1986, p. 628). As it stands, though, this question

⁵ See also Dummett (1992) and Dummett (1993).

seems too strong: in fact, Block thinks that we should weaken the requirement of individuating *identity* and *differences* of conceptual roles with the requirement of individuating a “multidimensional gradient of similarity of meaning”, so abandoning the “crude dichotomy of same/different meaning” (Ibid., p. 629). However, neither in that paper nor in following papers Block makes any point regarding this crucial question, and the recent story of the problem seems to have followed different paths.

In recent times many attempts have been set in order to tackle this difficulty. Some authors have declared that the notion of semantic holism cannot be maintained any more, committing themselves to weaker approaches to the relation between compositionality and the individuation of semantic values, arguing that the semantic elementary unit are to be found at the level of sentences. This view, called “molecularism”, has been particularly defended by Dummett (1992; 1993) and Perry (1994). This view stems from Wittgensteinian consideration regarding language games as complete languages applicable just to specific circumstances, and is at the origin of the idea that holism can be considered as true only in a local sense (e.g. Bilgrami 1986)⁷. A consequences of such a view is that the bad consequences of holism are the hallmark that we are in part epistemically incompetent on everything, so that whenever a new inference commit us to modifying the content of a thought this amount to our learning something new.

A development of such a view is in Artificial Intelligence, where narrowing down *meanings*, as general concepts, to *meaning in context* has proven to be useful in solving the problem of communication among speakers with different beliefs and thoughts. The idea is that contexts offer a selection of some roles in the individuation of contents (cf. Penco and Vignolo in press). While excluding some roles, contexts are not enough for determining exactly the constitutive roles of an expression. So, the crucial step is provided by the convergence of two or more speakers on partial aspects of the situation that is the subject matter on some inferences through the filtering of some details and by considering different perspective from which to consider the same situation (cf. Benerecetti, Bouquet and Ghidini 2000) resulting in a sort of conceptual blending among the different perspectives (cf. Guha and McCarthy 2003).

A very different line of approach has been pursued by Eric Lormand (1996). This author, who is mainly interested in the impact holism has on the “meaning of mental representations”, here I would say the content of mental representations, argues that holism threaten semantic stability so that “If meanings are holistically dependent, then semantic differences anyway seem

⁶ Another source of criticism on CRS has to be found in Searle’s arguments against artificial intelligence, in particular his Chinese Room argument (Searle 1980). For a comprehensive reply to this argument in relation to CRS see Rapaport (2002).

to balloon into semantic differences everywhere” (Ibid., p. 51). However, Lormand thinks that, given a specific mental content, there is no way to individuate a sharp or vague criterion to separate relevant from irrelevant contents or inferences that constitute it. So, he endorses a very radical holistic view, one according to which the content of a thought is determined by the contents of *all* thoughts in the system. In order to get to semantic stability Lormand proposes to block the inference that goes from “1) if a representation in a system S changes its meaning, all representations in S change meaning” to “2) If a representation in S changes meaning, no representation in S has the same meaning it had before the change” (Ibid., p. 57). In order to do so, the basic idea is to argue that each representation has more than one content. So, any mental representation concerning *a* constitutes the content of *a* itself, so that the content of *a* results from the collection of all the thoughts and inferences in which either *a* explicitly occurs or it is somewhat concerned. This entails that any mental representation has multiple contents simultaneously (he says, multiple meanings), an idea accepted in all its consequences, such as the following: “On the multiple meaning view, acceptance of a mental representation realizes multiple beliefs simultaneously” (Ibid., p. 58). For instance, the multiple meaning view amounts to arguing that the representation [cat] has as meanings representations such as [purring animal]; [the kind of thing mommy calls ‘cat’]; ... Consequently two individuals, or the same individual over time, can share the representation of [cat] because they share (at least) an overlapping non empty set that comprises at least one meaning. Semantic stability is then reached through the acceptance of a common overlapping meaning. However, as Kelly Becker has noticed (1998) this proposal seem to be caught in a circle, because the meanings of a mental representation are individuated through other mental representations, that necessarily calls for the same kind of instability blocking mechanism. But the blocking mechanism crucially presupposes the presence of an overlapping non empty set. Since there is no way to assure that in every situation the overlapping set will be non empty, this inflationary strategy seems doomed to failure.

Other strategies follow a more conceptual line. In a very interesting essay, Peter Pagin (1997) has noticed that if meaning holism is construed not as a way for *individuating* meanings, or contents of thoughts, but rather as a way of *determining* them – so recognizing that contents depend on the whole set of contents held true by an individual but are *not identical* with the whole set of inferences in which they participate - then global roles, that is the role that a given item has in all the inferences in which it participates, would determine meanings of contents without thereby individuate them. The crucial step is in the recognition that there is not a one-to-

⁷ But see Bilgrami (1988) for a development.

one relation between roles and contents; rather this is a many-to-one relations in that roles determine contents but not the other way around. In this way, one may change the inferences one is ready to assent to without this implying the changing of the contents of one's thoughts.

Greenberg and Harman (forthcoming) have argued along the same lines by showing that CRS, their version of the view here under consideration, is not committed to identify meaning with use or inferences but, rather, to the dependency of the former on the latter. This dependency relation is compatible with two people having different uses, or accepting different inferences, while sharing, at the same time, the same meaning. If you like, the dependency relation allows for multiple realizations so that the content of a thought, or the meaning of an expression, can be multiply realized in different mental inferences, or linguistic uses. In this way, contents supervene on roles.

This view tackles a broader issue concerning role's theories in general. Role's theories conceive meaning or content as a relational entity, not as a punctuate one. As such, relational entities can be, at most, similar to each other unless the entire system of relations to which they pertain is identical with some other system. So we should consider *similarity* of meaning or content instead of identity. Now, what are the criteria for similarity? When is it that some content (meaning) is similar to another? The notion of similarity cannot be defined in isolation. Something is similar to something else with respect to a common ground. For instance, my shirt could be similar to your in that both are made of cotton, but sizes, colours and shapes could be quite different. Because, as we know, these theories are crucially based on roles, content similarity is *determined* via role similarity. However, there seems to be no definitive or a priori answer on how deep or far we should consider conceptual and inferential roles to license a judgement of similarity. So, even if Pagin is correct in his general analysis, since there is no a priori way to solve this question *prima facie*, it is then advisable to turn our attention toward more empirical strategies. A possible strategy could be the following: the judgement of similarity is something that must be available and acceptable by actual speakers and thinkers. It is among those that use and understand meaningful and contentful items that such a strategy has to show its viability. It is in this perspective that I propose to take inspiration from the notion of *reflective equilibrium* as a method for indicating which are the roles that constitute the content of a mental item or the meaning of a linguistic one. It could be useful to spend some words introducing the notion of *reflective equilibrium*.

Reflective equilibrium and FRS

In general, reflective equilibrium is a state in which parts are harmonized in a coherent whole. The basic idea of this notion is to test various parts of a system against other parts so to have a whole consistent system. When the parts are in reciprocal agreement the condition of equilibrium is reached. More specifically, reflective equilibrium holds between abstract principles and actual practice or between general rules and particular applications of them, and the notion applies once that principles and practice are in reciprocal agreement. Nelson Goodman has introduced this idea with respect to the problem of justifying deductive and inductive inferences, and important extension of it have been proposed by Rawls (1971), Daniels (1996), Cohen (1981) and discussed by Stich (1990). On what basis, asks Goodman, may we say that an inference is valid? We cannot make appeal to some self-evident axioms nor to the very nature of human mind. Rather, he says:

Principles of deductive - and inductive - inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences we actually make and sanction ... This looks flagrantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend* (Goodman 1955, pp. 66-7).

The extension I am proposing has the aim of determining which functional roles are constitutive of content of thoughts. The property of *being constitutive* as applied to content of thoughts is so my counterpart of Goodman's property of *being justified* as applied to principles and practice of deductive and inductive inference. The intuition behind this parallelism is the following: according to Goodman justification entails acceptance, where acceptance is something that we show in our practice regarding rules and inferences. So to say, the game of accepting and rejecting⁸ is the shared ground on which any process of justification has to rely on for being considered. If there were no possibility of recognizing a form of acceptance or rejection then it would not be possible to consider the very notion of justification. Now, in our activity of individuating thoughts and understanding sentences something similar is the case. The notion of semantic constitutivity entails the possibility that every thinker or speaker be able of recognizing this property. However, this ability should not be considered as a perfect and

complete ability, that is, it is not necessary that each of us know in every detail what constitute what, but that at least has some grasp of what could be essential for someone to understand or grasp some concepts or thoughts. This means that if there were not semantic acceptance or rejection there would not be semantic constitutivity. Also in the case of the individuation of the content of thoughts or the meaning of sentences, then, we must presuppose a shared ground on which such activities can be performed. This shared ground is the possibility of recognizing semantic acceptance or rejection. Without this shared ground no semantic activity would ever be possible. We must start somewhere, and this is Goodman's starting point which I endorse. It should be noticed that this is not question begging: I am not arguing that semantic constitutivity is there from the beginning, but that semantic acceptance or rejection is, in parallel with Goodman's point.

Now, in the present context, to what are acceptance and rejection to be applied? So far we have been considering concepts and thoughts, and it is pointless to say that a concept is accepted or rejected if considered in isolation, and the same is true in case of thoughts. Consider the concept [cat]. As such, this concept cannot be judged to be properly or improperly used unless it brings toward, or is framed within, a complete representational state, such as a belief or a desire, assuming that these are the basic representational states. It is when the concept [cat] is enrolled in beliefs, such that *the cat is on the mat* or *that mountain is a cat*, that we can consider the acceptance or rejection of this concept's use, or of the others tied to it. The same can be said of thoughts. When we want to decide whether a thought is acceptable or not, we contrast it with other thoughts it may lead to, namely, with the inferences it is possible to draw from it. The difference between concepts and thoughts in this respect is sufficiently clear: the context of a concept is the thought in which it may be framed or bring toward; the context of a thought are the inferences, namely other thoughts, it may bring toward or is framed within. Acceptance and rejection, then, apply to two categories of semantic entities: thoughts, such as belief and desires (for short we might consider just beliefs) and inferences.

It is now time to set the present use of reflective equilibrium with respect to Goodman's original notion. A dissimilarity has to be faced: while Goodman had general principles and actual practices what do we have? We are considering just roles, so where are the abstract principles and where the actual practices? The only elements we have been considering so far are conceptual roles, as used in endorsing beliefs and desires, and inferential roles, as used in regulating processes of thoughts. I think we have to be happy with these elements, and we

⁸ Where rejecting has to be construed as the first step in the amending process.

should consider them as those that enter in the process of reflective equilibrium. The key idea is that these elements, once in reflective equilibrium, can give us what constitutes the content of a thought or the meaning of a sentence.

Before saying in what way these elements enter in reflective equilibrium, we have to consider another pair of items that should be contrasted in our version of this notion. These are, on the one side, the conceptual roles as used in beliefs and the inferential roles as used in thoughts of ourselves as individuals and, on the other side, the roles, both of conceptual and of inferential nature, of other individuals, these other people considered both as actual fellows and abstract individuals. That is to say, acceptance or rejection is a public phenomenon, then also other people have to enter the stage. These two pairs of elements, conceptual and inferential roles of ourselves on the one side and conceptual and inferential roles of other people on the other, should be coupled and crossed in the following way. When we consider our conceptual roles, that fix the beliefs we endorse, we should consider other people's inferential roles and when we take under scrutiny their conceptual roles we should consider our inferential roles and vice versa. It is the reciprocal agreement, or disagreement, that determines whether a conceptual role is constitutive of the content of a belief or whether an inferential role is constitutive of the content of a thought. To present this idea in parallel with Goodman's statement we can say: the conceptual and inferential roles that are constitutive of the content of a belief or of a thought are those that meet the condition of reflective equilibrium. To complete the parallelism all the way down we should consider that conceptual roles fix the beliefs we endorse and inferential roles the thoughts we entertain. Consequently we can consider beliefs and thoughts as the proper elements for the reflective equilibrium, being those the ones that are recognized in our common practice of semantic acceptance and rejection. It is now possible to set the notion of reflective equilibrium in details. It goes as follows:

Other people inferences are amended if they yield to beliefs we are unwilling to accept;
Other people beliefs are rejected if they violate inferences we are unwilling to amend.

The inferences and beliefs that are neither amended nor rejected in a public confrontation are *constitutive* of the content of a concept or of a thought. Analogously, these beliefs and inferences may be said to be constitutive of the meaning of a word or of a sentence.

Before considering a couple of examples, a general clarification is needed. I am taking FRS in the context of an interpretative theory of thought's ascription⁹. This move is at odd with, for instance, Block's construal of FRS. Block thinks that FRS applies directly to a system of internal representations, a language of thought. However, de Saussure (1922), that can be considered the father of this theory (see Lepore 1994, p. 193), took it applying to public linguistic manifestations. It seems to me, then, that endorsing FRS does not necessarily imply endorsing the language of thought hypothesis, and we can be neutral with respect to this issue. On the contrary, it is admissible that the public manifestation of our thoughts can be understood in terms of FRS, regardless whether this public manifestation stands on its own or it is the result of the computational activity of a language of thought. That is, we can consider our thoughts and beliefs, as they are expressed in sentences or ascribed through the interpretation of behaviour, as being in functional relations to one another, and in the same way individuated, an idea already expressed by Sellars (1963).

It is now time to go back to the definition of reflective equilibrium as applied in the present context and test it against an example, considered both in the positive and in the negative. Take Putnam's case of "this is water" said by Ancient Greeks when looking at one river. If we exclude referential considerations, how can we be sure that we are using "water" in the same sense Ancient Greeks did? Even if, as Block urges, we were not considering content identity but content similarity, how would we determine the latter? I argue that similarity can be determined by looking at those beliefs of the Ancient Greeks that do not violate the inferences we are unwilling to amend and those inferences of them that do not yield to beliefs we are unwilling to accept. These beliefs and inferences are what we and the Ancient Greeks share with respect to the concept [water] and to the content *this is water*. Those are the beliefs and inferences that are *constitutive* of the content of any thought or concept concerning water. So, for instance, suppose that the thought *this is water* brings inferentially to the thought *water flows*. Because the inference does not violate any belief we have concerning water, this inferential role can be considered constitutive. The same applies for an inference of this sort: after having seen many rivers and having said in each case "This is water" the Ancient Greeks would infer "In all rivers there is water".

Let me now explore the example in the negative sense. What about their belief *Water is one of the four simple elements out of which the entire universe is composed*? This belief violates many inferences we are unwilling to amend, such as *if there is water then there is a composed*

⁹ As those espoused by Davidson and Dennett.

entity. In this case, we would say that the belief of the Ancient Greek is *not* constitutive of the concept of water. And what about the inference of the Ancient Greeks *If there is a river, then there is a God*? This would yield a belief we are unwilling to accept, such as *For every river there is a divinity*. In this case, we should be ready to abandon these inferences and beliefs, rejecting them or asking, so to say, the Ancient Greeks to revise them. In a word, these are *not* constitutive of the concept of river.¹⁰

Generalizing this point, we can say that the beliefs and the inferences that are constitutive of our concepts and thoughts are those that are shared and have been shared in the sense given by the above-mentioned construal of reflective equilibrium. It is this subset of all beliefs and inferences with respect to any concept or thoughts that is constitutive of any content. What is not shared is part of the culture and the knowledge of the time or of a certain tradition, but it is not constitutive of the content or meaning. From this follows that an individual or a society with which we do not share anything, not even the roles concerning logical constants, would be radically unintelligible. Through this method, then, we can individuate those beliefs and inferences that are constitutive of the content of our thoughts, or of the meaning of our sentences, without invoking obscure claims or the analytic/synthetic distinction.

An immediate objection could be the following: it is not the case that we do not accept the belief that *water is one of the four elements out of which the entire universe is composed* because it yields to an inference we are unwilling to amend. Rather, we reject that belief because it does not agree with all the other beliefs we have concerning physics and chemistry. This objection, however, assumes that what science says is not entrenched in our linguistic practice. As if science were out of our language. Moreover, as it is the objection is not in the spirit of FRS. FRS is the claim that the content of a concept or of a thought is given by its role in the overall pattern of beliefs and inferences. This means that, for instance, it would be impossible to determine the content of a belief in isolation. When we consider the acceptability of a belief or of an inference we are judging how it fits in the other beliefs and inferences we have. If the belief considered would yield too many transformations in our conceptual and inferential net, in particular transformations that imply either that we have to accept or to amend something we are unwilling to do, then this belief is rejected. That is why we have to test other people's beliefs and inferences against the epistemic states and processes we endorse.

Another possible objection may use the Twin Earth example. My twin and me are disposed to have the same beliefs and to draw the same inferences. Would it follow that we would have

¹⁰ Because we cannot amend their inferences, we simply abandon the premises of the inferences.

the same content? First of all, it should be noticed that the application of reflective equilibrium to FRS is particularly devoted to narrow content, that is, content individualistically considered. But many FR theorists, for instance Block and Harman, have considered this kind of content as relative to contexts. In this way, the global content of a thought is given by the narrow content in a context, context that provides the wide content. So, twins would share the narrow content but not the wide content, hence not sharing content globally considered.

A final consideration regarding this hypothesis on how to solve the main difficulties concerning FRS. The idea of considering other people beliefs and inferences is what constitute a practice well entrenched in our cognitive life, the so called *folk psychology*. Through it we are able to predict and explain other people's behaviours. In recent time, many cognitive scientists have thought that this practice has a mental counterpart in a specific module which would implement a sort of *theory of mind*. It is this theory that governs the functioning of the module making explicit our capacity of being folk psychologists¹¹. Reflective equilibrium, then, could be implemented by a theory of mind, giving to this argument not only a metaphysical aspect concerning the condition of individuation for concepts and thoughts but also an empirical and epistemological twist. It gives the identity conditions according to which we may say whether two concepts or thoughts are the same or not, and which gradient of difference there could be, and at the same time offers a way for understanding how such individuation condition could be rooted in our cognitive capacities. The fact that such a view integrates the original idea of Sellars of thoughts as the result of each others' interpretation with recent theories of psychology should not come as a surprise, as other (e.g. Rosenberg 2004) have shown.

Summing up, I offered a specific way out from some of the troubles that affect FRS. The most important feature of this approach is that it gives us indications on *how* to select the beliefs and inferences that are constitutive of the contents of our thoughts. In this sense, an extension of this approach to the problem of meaning would be direct and this, I think, is another advantage of the view.

Acknowledgements

A previous version of this paper has been read at the European Conference of Analytic Philosophy, held in Maribor (in 1999). I wish to thank Carlo Penco, Eva Picardi, Silvano Tagliagambe, Alberto Voltolini and an anonymous referee for suggestions on a previous version.

¹¹ The literature on this topic is by now abundant. It has to do both with the general problem of the theory of mind module and the debate on how this module operate, that is, either through a sort of simulation or by means of a general theory. On these topics see Premack and Woodruff (1978) for the original hypothesis, Carruthers and Smith (Eds) (1996) and T. Stone and M. Davies (Eds) (1995) for two surveys.

References

- Becker, K. (1998) "On the Perfectly General Nature of Instability in Meaning Holism" *The Journal of Philosophy*, 95, pp. 635-40.
- Benerecetti M., Bouquet P. and Ghidini C. (2000) "Contextual Reasoning Distilled", *Journal of Theoretical and Experimental Artificial Intelligence*, 12(3), pp. 279-305.
- Bilgrami, A. (1986) *Belief and Meaning*, Blackwell, Oxford.
- Bilgrami, A. (1998) "Why Holism is Harmless and Necessary", *Philosophical Perspectives*, 12, pp. 105-26.
- Block, N. (1986), "Advertisement for a Semantics for Psychology", in P. French, T. Uehling e H. Wettstein (Eds), *Midwest Studies in Philosophy*, Vol. X, University of Minnesota Press, Minneapolis, pp. 615-78.
- Block, N. (1993), "Holism, Hyper-analyticity and Hyper-compositionality", *Mind and Language*, 8, pp. 1-26.
- Block, N. (1998a) "Semantics, Conceptual Role" *The Routledge Encyclopedia of Philosophy*,
- Block, N. (1998b) "Holism, Mental and Semantic" *The Routledge Encyclopedia of Philosophy*,
- Block, N. (1999), "Functional Role Semantics" in R. Wilson, F. Keil (Eds), *The Mit Encyclopedia of the Cognitive Sciences*, Cambridge Mass. Mit Press, pp. 331-2.
- Carruthers, P. Smith, P. (1996) (Eds), *Theories of theories of mind*, Cambridge University Press, Cambridge.
- Churchland, P. M. (1998) "Conceptual Similarity across Sensory and Neural Diversity: The Fodor/Lepore Challenge Answered", *The Journal of Philosophy*, 95, pp. 5-32.
- Cohen, J.L. (1981), "Can human irrationality be experimentally demonstrated?", *Behavioral and Brain Sciences*, 4, pp. 317-70.
- Cohen, J. (1999) "Holism: some reasons for buyers' remorse", *Analysis*, 59, pp. 63-71.
- Daniels, N. (1996) *Justice and Justification*, Cambridge University Press, Cambridge.
- Dummett, M. (1973) *Frege, Philosophy of Language*, Duckworth, London
- Dummett, M. (1992) *The Logical Basis of Metaphysics*, Duckworth, London.
- Dummett, M. (1993) *The Seas of Language*, Clarendon, Oxford.
- Field, H. (2001) "Attribution of meaning and content" in H. Field, *Truth and the Absence of Facts*, Clarendon Press, Oxford, pp. 157-74.
- Fodor, J.A. (1987), *Psychosemantics*, Mit Press, Cambridge Ma.
- Fodor, J.A. Lepore, E. (1991), "Why Meaning (Probably) Isn't Conceptual Role", *Mind & Language*, 6, pp. 328-43.
- Fodor, J.A. Lepore, E. (1992), *Holism*, Basil Blackwell, Oxford.
- Fodor, J.A. Lepore, E. (1993), "Reply to Block and Boghossian", *Mind & Language*, 8, pp. 41-48.
- Fodor, J.A. Lepore, E. (2002), *The Compositionality Papers*, Clarendon, Oxford.
- Greenberg, M. Harman, G. (forthcoming) "Conceptual role semantics", in E. Lepore and B. Smith (Eds), *The Oxford Handbook of the Philosophy of Language*, Oxford.
- Guha R. McCarthy J. (2003), "Varieties of contexts", *Lecture Notes in Artificial Intelligence – CONTEXT 2003*, pp. 164-177.
- Harman, G. (1987), "(Nonsolipsistic) Conceptual role semantics", in E. Lepore (Ed), *New Directions in Semantics*, Academic Press, London, pp. 55-81.
- Heal J. (1994) "Semantic holism: still a good buy", *Proceedings of the Aristotelian Society*, 94, pp. 325-39.
- Heal J. (1999) "Thoughts and holism: reply to Cohen", *Analysis*, 59, pp. 71-78.
- Lepore, E. (1994) "Conceptual Role Semantics" in S. Guttenplan (Ed) *A Companion to the Philosophy of Mind*, Blackwell, Oxford, pp. 193-200.
- Lormand, E. (1996) "How to Be a Meaning Holist", *The Journal of Philosophy*, 93, pp. 51-73.

- Pagin, P. (1997) "Is Compositionality Compatible with Holism?" *Mind & Language*, 12, pp. 11-33.
- Penco, C. Vignolo, M. (in press) "Converging toward what? Pragmatic and Semantic Competence" in P. Bouquet and L. Serafini (Eds) *Context representation and reasoning 2005, Proceedings of the first international workshop* (CEUR workshop proceedings) 2005.
- Perry, J. (1994) "Fodor and Lepore on Holism", *Philosophical Studies*, 73, pp.123-58.
- Premack, D. e Woodruff, G. (1978), "Does the Chimpanzee Have a Theory of Mind?", *Behavioral and Brain Sciences*, 1, pp. 515-26.
- Putnam, H. (1988), *Representation and Reality*, Mit Press, Cambridge Ma.
- Putnam, H. (1990) *Realism with a Human Face*, Harvard University Press, Cambridge Ma.
- Rapaport, W. (2002) "Holism, Conceptual Role Semantics, and Syntactic Semantics", *Mind & Machines*, 12, pp. 3-59.
- Rawls, J. (1971) *A Theory of Justice*, Harvard University Press, Cambridge Ma.
- Rosenberg, J. (2004) "Ryleans and Outlookers: Wilfrid Sellars on 'Mental States'", *Midwest Studies in Philosophy*, 28, pp. 239-65.
- Saussure, F. de (1922) *Cours de linguistique générale* (Ed. by C. Bally and A. Sechehaye), Paris, Payot; tr. and annotated by R. Harris, *Course in General Linguistics*, Duckworth, London, 1983.
- Sellars, W. (1956), "Empiricism and the Philosophy of Mind", in H. Feigl et al. (Eds.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, University of Minnesota Press, Minneapolis.
- Stich, S. (1990) *The Fragmentation of Reason*, MIT Press, Cambridge Ma.
- Stone, T. Davies, M. (1995) (Eds), *Mental simulation: evaluation and applications*, Blackwell, Oxford.