How Are Semantic Metarepresentations Built and Processed?

Manuel Bremer

Abstract

This paper looks at some aspects of semantic metarepresentation. It is mostly concerned with questions more formal, concerning the representation format in semantic metarepresentations, and the way they are processed.

Section 1 distinguishes between metacognition and metarepresentation in a narrow and broad sense. Section 2 reminds the reader of some main areas where metarepresentations have to be used. The main part considers the ways that metarepresentations are built and processed. Section 3 introduces some general ideas how semantic metarepresentations are built and processed. Section 4 looks at some recent theories about ways that semantic metarepresentations are built and processed.

1 Introduction

In psychology *metacognition* is described in its function and in a way that corresponds to ways it can be tested and measured (cf. [9]). Psychology here works with the concept of metacognition in its bare sense of a cognition being the object of an other cognition. Some theories propose models of control flow and access between cognitive faculties (like long term memory and linguistic reports on stored information) or 'files' of stored information and ongoing cognitive processes. These processes as cognitive processes need not be conscious, but as testing relies on linguistic reports or some form of explicit judgement (e.g. by making a mark on a confidence level scale) conscious forms of metacognition or at least the conscious states correlated with metacognitive states (like feelings of knowledge) are in the foreground of laboratory tested psychological models.

Kriterion – Journal of Philosophy (2012) 26: 1–17.
http://www.kriterion.at

What, peculiarly, is missing in these theories and observations are fine-grained models of the format(s) and *representational* structure(s) of these meta-cognitions. Do we deal here with a linguistic/verbal interpretation of feelings of security (of knowledge, say)? Or do we deal here with a *language of thought*-representation in the scope of another *language of thought*-representation, which by the usual mechanisms of belief report then becomes verbalized? Or something else altogether...

Apart from the representational structure one may wonder about the access mechanisms involved in these forms of meta-cognition. Are memory-stores accessed? And if so, to what depth? Might there be a relation between depth of search (in a file or store) and the confidence judgement concerning the obtained result?

Ultimately the observations like confidence judgements being influenced by frequency of presentation of material (notwithstanding proper temporal order, as in witness accounts) should have an explanation in terms of a model of arriving at such judgements by a control flow accessing different (memory) stores and faculties.

A meta-representation in the *narrow* sense is a representation the content of which contains another representation. Quotations, codings (like Gödel-numberings) or higher order beliefs are taken to be typical examples.

Representations concerning our cognitive (representational) faculties may be taken as meta-representations in a *broader* sense. They do not contain individual other representations, but their content contains or refers to representational properties (i.e. either properties of individual representations in their function as representations or properties of some faculty inasmuch as they are invoked in the explanation of its representational function). A justification of a claim may invoke other representations or may invoke beliefs about the proper workings of claims of this type (e.g. beliefs about the reliability of observation). In this case the justification is metarepresentational in the broader sense.

A study of meta-representation belongs in part to the study of cognition (cognitive science) or the study of the mind in the narrow sense of exploring how the mind works and where in its workings metarepresentations occur, and how they function. A study of metarepresentation also belongs to epistemology, and thus the theory of cognition is a broader sense, in as much as meta-representations have a role to play in justification, a theory of justification thus containing a part that deals with the role and function of metarepresentations.

2 *Why Are Meta-Representations Built?*

In the cognitive sciences metarepresentations are investigated from several ankles. Theories invoking a crucial role for metarepresentations range from higher order theories of consciousness, theories of self-monitoring, theories of attitude attribution to (pragmatic) theories of communication and implicatures. An impression of the breadth of studies can be gained from the collection *Metarepresentation* ([18]). Even in the wider public theories of folk psychology as (innate) 'theory of mind' and studies on the metacognitive abilities of animals have received wide interest. I will focus here instead on two areas which are closely related to semantics and to our concern of the structure of metarepresentations.

One area in which metarepresentations are crucial are theories of justification, especially coherence theories of justification. Judgements of coherence (of one belief/statement cohering with others) are metarepresentational. Coherence theories, thus, often work with the idea of the reasonser accessing the belief systems and being able to (meta-)represent beliefs. Laurence BonJour ([3]) calls this 'the Doxastic Presumption': the believer 'must somehow have an adequate grasp of the total system of beliefs, since it is coherence with the system which is at issue.' (p. 102). The 'grasp' consists in 'metabeliefs' (operating mostly as presumptions that justificatory beliefs are in store). Keith Lehrer's subjective coherentism (cf. [15]) is another theory involving metarepresentation of this kind: 'evaluation or certification of incoming information is a metamental activity. The mind that certifies incoming information is a kind of metamind' (p. 252).

Theories of belief update and justicatory relations are linked to semantic theories, as the question which information or newly believed sentence affects which other information or sentences is related to the meaning and content of the sentences in question. In semantics itself the question of metarepresentation arises in several connections.

Speakers of natural language come by with limited expertise on employing terms like "beech" or "lime (tree)" facing a tree. Linguistic division of labour allows that we defer to the experts. In this case we employ a term with the knowledge that this very term has appropriate conditions of usage and proper reference, only we do not (exactly) know which. This linguistic knowledge again is metarepresentational because it has to quote the term it is knowledge about. If I know 'The term "beech" refers to a tree identifiable by botanic experts' I have a metarepresentation concerning the lexical item "beech".

Speakers also have to have some accessible, though often sub-doxastic-

ally used, knowledge of the semantic rules of their language. Updating one's description of the world in case of conflicting data or expectations includes metarepresentations concerning proper usage. To rehearse the basic idea: According to the model of radical translation (cf. [8]): We translate the statements of L_1 into statements of L_2 which give the truth conditions for L_1. To do this we look at the linguistic behaviour of the speakers of L_1. An interpreter proceeds by correlating the statements to be interpreted with the situational conditions he perceives (i.e. with his perceptions and not with his physiological states). The reference to situations of justified usage (where the truth conditions of some L_1 statement are met) enables the interpreter to formulate an interpretation axiom leading to (T)-equivalences in the truth/meaning theory for that language. To accomplish this the interpreter incorporates normative assumptions with regard to the L_1-speakers. To start with we transfer our logic to L_1. Secondly we have to assume awareness of propositional attitudes. For something to be a reason the reasoner must be aware of it, or at least he could bring it to his awareness. We interpret by assuming that the statements build a coherent system. Without these assumptions understanding would be impossible. Someone who would use expressions arbitrarily would make it impossible to establish a correlation between his manners of usage and situations in the world. If, on the other hand, the use of expressions builds a coherent system, then statements which are supposed to be true will be integrated in the belief system, and statements which turn out to be false will be taken out. To do this speakers have to have propositional attitudes. They believe that something is the case, and believe that there are connections between what they believe (inferential relations between statements). And they believe that there are rules determining how the expressions of L_1 should be employed. For example that some new circumstances no longer allow to speak of an object a being F, since under the new conditions "F" should not be employed. To judge the coherence of speaking an interpreter has to know what should be said in L_1 under some circumstances. By this the interpreter has understood the assignment of truth conditions as normative. One has or formulates a theory of meaning for L_1 with the maxim: Use the expressions of L_1 under exactly those conditions which are specified in the (T)-equivalences (or meaning postulates).

Any theory of reading off the coding of concepts with words of some (natural) language by radical interpretation or some related method commits itself, therefore, to the existence, accessibility, and constitutive force of semantic rules in that language.

3 How Are Meta-Representations Built and Processed?

As a point of departure let us assume some version of a *representational theory of mind* (RTM) where beliefs are understood as tokens of *language of thought* (LoT) sentences processed by some cognitive (sub-)system or stored in a 'belief box'. A dispositional or sub-doxastic belief needs nothing besides that LoT-sentence. An occurring belief if it is accessible to consciousness involves some further representation (like a natural language sentence verbalized in inner speech) as LoT-sentences are neither phenomenally given nor immediately accessible as such. Even some sub-doxastic or dispositional beliefs may be tied to some specific way to express that belief (by mechanisms of memory or by limited expressive power of the cognitive system under discussion). If it is true that John dispositionally believes that the Earth is flat, John may have never entertained the sentence 'The Earth is flat' or utter it on being asked. If it is true that John has an occurring belief (in inner speech) that the Earth is flat, he has at least used some natural language sentence, synonymous to 'The Earth is flat".

In a representational theory of mind there are several levels of mental processing, some of them are representational, some of them are subsymbolic, some representations we are aware of, others we are not aware of. Functional architecture comprises levels of intentionality, levels at which stimuli are transduced into representations to be processed at some intentional level, and ultimately some physiological implementation level.

Metarepresentations are vital in *de dicto* attitude attributions. A short reminder may be in place.

Every natural language belief report has a *de dicto* and a *de re* reading in semantics or logical form, the surface sentence thus being semantically ambiguous. The two readings differ in truth conditions. This does not only pertain to singular terms but to all constituents with a semantic role.

(1) John believes that the gardener is at sleep.

has a *de dicto* reading

(1d) Believes(John, The gardener is at sleep)

In the *de dicto* reading John stands in the BELIEVE relation to a sentence either identical or at least synonymous (identical in meaning) to the sentence used (not mentioned) in the "that"-clause.[1] What we understand as listeners to the report is understood by John.

(1) also has a *de re* reading, using some form of propositional/sentential quantification (supposedly with a substitutional semantics):

(1r) $(\exists p)((p \equiv$The gardener is at sleep$)\wedge($Believes(John,$p)))$

In the *de re* reading the reporter claims that John's belief has some objective *content*, however John referred to that content. The sentence used in the report need not share its meaning with the sentence/proposition John believed. It only shares its referential content. *De re* attribution may concern dispositional or sub-doxastic beliefs. *De dicto* attributions presuppose that the attributee has or has had some occurent – maybe even consciously accessible – belief that p.

Given (1r) not only singular terms but any part of the sentence is open to extensional substitution (say by some other phrase picking out the processes going on in the gardener when being asleep).

The *de re* attribution is true iff John stands in the BELIEVE relation to some sentence equivalent to that used in the "that"-clause. Thus is true if John tokens or stores a LoT-representation (in his 'belief box') having the same content as the sentence used in the "that"-clause. The *de re* attribution, however, can also be true, because John has only the dispositional belief that p. The *de dicto* belief attribution is true iff John has or has had an occurring belief *using* an introspectively given – typically verbalized – representation with the same meaning as the sentence used in the "that"-clause. *De dicto* reports thus are essentially metarepresentational. This need not be so for *de re* reports: John may have the dispositional believe that New York is not in the Netherlands, although he has never explicitly thought about it, since it is implied by some of his explicit beliefs about New York. As John has never had any representation processed equivalent to 'New York is not in the Netherlands' one may doubt such a *de re* attribution to be metarepresentational. The person attributing the dispositional belief to John uses her own representational resources and need not even aim at claiming anything about John's representations.

Belief attributions expressed in natural language are – one may think 'easily' – metarepresentational by quoting another sentence or using that other sentence in an embedded complement clause. Our sub-doxastic reasoning, however, will use such attributions as well. And the natural language reports have to have some conceptual content. The representation medium of these levels (the LoT) therefore has to have the means not only to built metarepresentations in general, but to built metarepresentations which contain items of the public language.

There are several ways in which meta-linguistic representations may occur in the LoT:

a) Linguistic sharing of labour may be a way to acquire a concept THAT-WHAT-IS-REFERRED-TO-BY-EXPERTS-AS-α, α being a structural description (say a quote of a form in one's linguistic community). In this case, supposing a successful hooking up to the target extension, at least the mediation between the new concept (a LoT-type) and the referent requires meta-linguistic representation. Some time later that concept may be linked to a as its expression. One may then have forgotten who the experts are/were. One knows, nevertheless, that a is an established lexical item of one's language, which expresses the concept (formerly known as) THAT-WHAT-IS-REFERRED-TO-BY-EXPERTS-AS-α. The concept THAT-WHAT-IS-REFERRED-TO-BY-EXPERTS-AS-α refers to α's referent. *For experts* this is the reference of the non meta-linguistic $concept_A$ they express by α.

b) Intensional contexts invite attributions of propositional attitudes which essentially point to *the way* the attributee represents a state of affairs. In such attributions one may meta-linguistically point to a speaker's idiolect, quoting an expression of the language to explain its usage by the attributee of the attitude ascription. The conceptual content of such an attribution thus contains a quotation or some other meta-linguistic device (like reference to phonetic or graphemic features). [see section 4.1; cf. already [14]]

c) Rules of grammar allow for stylistic variations. Such variations may be the vehicles of language shift (cf. [1, pp.218–223]). Such stylistic variations may be triggered or invited by lexical items (in the context). In this case the grammatical competence of a speaker contains rules which refer to other *lexical* items (i.e. they are meta-linguistic).

If a LoT representation contains or quotes a lexeme (a word) of a natural language, how is that lexeme to be represented? Certainly the mind need not process a sound file or a picture of a written word at that time. Syntactic derivations arrive at structural descriptions. Such structural descriptions are pairs of representations, one to be passed to the phonetic component, one to be passed to the conceptual system. Each part is a LoT representation. The representation π to be passed to the phonetic component contains all phonetic features needed for Spell Out[2]. So a word or part of a phrase is represented at a LoT level as a set of phonetic features, each of which has some LoT representation. Thus it

is sufficient for quoting a natural language word or phrase to embed its representation π into another LoT representation.

One usage of natural language representations embedded in LoT-processing may be *as labels* in file semantics. File semantics works with the idea that our knowledge is heavily compartmentalized. One compartment may contain my botanic knowledge about elm trees, another my knowledge about Cicero – and maybe another my knowledge about Tully. This solves a couple of problems: facts about the same object can be kept apart if they are filed in different places; keeping relevant facts from interacting may be an explanation of self-deception (cf. [7]). Merging files may be the use of informative identity statements (cf. section 4.1).

4 *Some Recent Proposals on the Structure of Semantic Metarepresentations*

This paragraph looks at some recent theories involving the use of metarepresentations in our semantic capacities. The first two sub-paragraphs discuss two theories crucially exploiting the presence of metarepresentations. The third sub-paragraph criticizes some neglect of metarepresentation in two other theories.

4.1 Fiengo and May on Belief Attribution and Informative Identity Statements

De Lingua Belief (cf. [10]) by Robert May and Robert Fiengo uses the proper name problems (like substitution, co-reference...) to illustrate their theory of meta-linguistic beliefs. Meta-linguistic beliefs are said to occur at the linguistic derivational level of Logical Form (and thus propositional content), and attributing such beliefs is said to improve accounting for language use (*inter alia* with respect to proper names).

The two central ideas of the book are (i) a distinction between names and 'expressions', which embed names, and (ii) an analysis of the logical form of some sentences which proposes meta-linguistic additional content (beyond presumable surface content).

(ad i) Fiengo and May claim that names 'do not refer' at all (p.14)! They are employed in 'expressions'. An 'expression' is a phrase using some phonological form carrying an index to distinguish it from another 'expression' using the same phonological form, e.g. $[\text{Fred}_1]$ vs. $[\text{Fred}_2]$. The co-indexing device can also be used to explain anaphoric reference (use of pronouns). Co-indexing thus does not require identity of

used phonological form. Referential knowledge consists in knowing 'assignments' which correlate 'expressions' with their referents. If 'expressions' are part of the logical form of a statement, seemingly tautological statements can be informative: [Paderewski$_1$] is [Paderewski$_2$]. And the logical form reveals the information in informative identity statements: [Cicero$_1$] is [Tully$_1$].

(ad ii) Fiengo and May distinguish *de dicto* attributions, as these include a commitment to the way the described person uses expressions, as involving meta-linguistic content from, ordinary, *de re* attributions. For instance: "Fred believes Cicero is a Roman" is taken as "Fred believes [[[Cicero$_1$] is a Roman] and ["[Cicero$_1$]" refers to Cicero]]. The last occurrence of "Cicero" may be exchanged by any other way to pick out the reference of the 'expression' [Cicero$_1$], e.g. to account for Fred mistaking somebody else for Cicero. Further on, the failure of substitution into *de dicto* attribution can now be explained. As the 'expression' is quoted in the second conjunct substitution would be substitution into quotation marks, which is forbidden.

These main idea account for the problems in the vicinity of proper name semantics, but if true they substantially revise our picture of semantics (e.g. which items refer, the role of the lexicon, the theory of the linguistic-conceptual interface)! As names are only used in 'expressions' one needs supposedly less lexical entries for the same phonological form, but this is no representational gain, as the authors complaining about the 'many names of the same phonological shape' (p.146) seem to think, as one needs now as 'many' assignment statements (for each possible referent of an 'expression' built by using that name). These assignments supposedly work as semantic axioms to be used in deriving truth conditions in internal semantics – and so forth.

Both main ideas invite a couple of questions. The indices which come with 'expressions' we find neither in verbal communication nor in written texts, nor in inner speech (verbal imagination). Thus 'expressions' occur on some sub-doxastic cognitive level, say of processing of syntax somewhere in the linguistic derivational system. One may ask now whether what carries such indexed labels are syntactic entities or concepts themselves. If one adheres to some LoT hypothesis one may say that a speaker may have two Paderewski concepts, each of which labels some memory folder with corresponding beliefs. The indices distinguish these concepts and the concepts are linked to the ambiguous name. If one learns about their co-referentiality the two files are merged. As the sub-doxastic 'expressions' have to be language of thought items

themselves the only reason to introduce these additional representations (beyond concepts and lexical items) can be some derivational advantage (like better anaphoric reference). In case of an ambiguous spoken name the standard picture sees the ambiguity arising at the level of interpretation (assigning a concept); Fiengo and May will see the ambiguity arising in derivation (building a different 'expression'). They need a new account of parsing/de-coding by the audience then as well. Phonological forms especially if mentioned in meta-linguistic beliefs have to be processed early in derivations now. One therefore has to reconsider basic derivational procedures (like Merge in Generative Grammar) – quite an agenda, it seems.

The second main idea (the proposal of meta-linguistic propositional content) invites similar questions. A whole conjunct present in Logical Form does not get spelled out (is not pronounced). One needs principles beyond those currently used in Generative Grammar to explain what governs Spell Out now.

4.2 Cappelen and LePore on Raising in Mixed Quotations

In their recent book *Language Turned On Itself* (cf. [6]) Herman Cappelen and Ernie LePore revise their former theory of quotation. Their new 'Minimal Theory' centres around the principle that the basic semantic rule for quotation is: ' '*e*' ' quotes '*e*' (for any expression '*e*'). So any quotation which quotes a quotable item contains that very item. This goes against theories (including their own former theory) which analyse so called 'mixed quotations' (like: Fred believed that 'the gardener' is sleeping) as *both* quoting and using the quoted item (in this case "the gardener"). Such theories appeal to the supposedly obvious fact that we understand the whole belief attributed: there are no black holes in understanding where a quoted item occurs. The attribution supposedly was ungrammatical if the quoted items had not their usual grammatical roles (in the example: "the gardener" as noun phrase forcing agreement with the verb phrase). As these theories have many difficulties – I do not go into here – Cappelen and LePore switch to the Minimal Theory. They have to deal now, however, with our intuitive understanding of mixed quotes. They develop a theory of raising the quote in Logical Form, a theory which – apart from being about quotation – is metarepresentational. The raising rules are metarepresentational. Given a mixed quotation (often a belief attribution like in our "the gardener" example) the raising rules proceed as follows (cf. pp.138-41): Out of the complement clause ('that the gardener is sleeping') the quoted item – usually

a phrase – is raised in the syntax tree to a new sister node of the complement leaving behind a trace. The semantic value of that new sister node is the quotable item, it is quoted as the sister node is a quotational phrase 'QXP'. The trace still points to the original phrase and is interpreted straight forwardly; in our example: the noun phrase 'the gardener' is raised out of the complement into a quotational phrase QNP which combines the functional head Q (which maps quotable items to their quotes) with it to yield "the gardener", the trace is interpreted as referring to the gardener. In short: Raising to quotational phrases generates mixed quotes, we understand the proposition in the mixed quote, since the trace still points to the former constituent.

The theory nicely explains *how* the metarepresentation is built. It invokes a quotational capacity (maybe some function in the LoT) by which we can embed quotable items into other representations without losing their objective content in the compositional content of that longer representation. There are many reasons why we are interested in quoting. And the theory nicely accounts for features of mixed quotation like reflexivity, which requires that in 'Fred said that "the gardener" loves herself' the quoted item c-commands the pronoun, which comes out true in the raising tree.

As Fiengo and May's theory the theory employs the idea of transformations in Logical Form, and it also affects Spell Out. In distinction to Fiengo and May's theory the concept of 'trace' (or 'copy') explains that at the Phonological Interface only the first occurrence is spelled out. Usually in Generative Grammar, however, transformations in Logical Form are considered to take place *after* Spell Out. Both theories thus have to be related to the latest development within the Minimalist Program, which completely rework Logical Form and Spell Out towards 'phases' (cf. [13]).

4.3 Patterson and Hanna on Sub-doxastic Logic and Semantics

In *Rationality and Logic* (cf. [12]) Robert Hanna tries to defend and re-vitalize the more or the less Kantian thesis that all rational human beings share a faculty of logic, which is governed by normative principles. This faculty is protological is the sense that it is operative in constructing logical systems. It is *a priori* by being innate. Making use of both the innateness idea with respect to cognitive faculties and of the idea of constructing individual logical systems by an innate faculty of principles Hanna takes up main elements of current cognitive linguistics and tries to combine them with ideas of the (Neo-)Kantian tradition. Just

as universal grammar is understood in generative grammar as a set of principles by which individual languages are learned (i.e. by specifying some parameters occurring in them), so Hanna conceives of the protological faculty. Individual logics are then a collection of separated systems which nevertheless share abstract common features, just as the collection of all human possible languages shares the common features of universal grammar. The Kantian mentalistic talk about faculties of reason is transformed into the cognitive science idea of innate capacities, which are modules of the mind/brain. Hanna thus brings together traditional mentalism with mentalism in the tradition of Jerry Fodor and Noam Chomsky. Hanna tries to integrate the idea that humans are essentially rational beings with current theories of cognition and the modular mind. This in itself is a fruitful approach, since it either provides us with the opportunity to integrate traditional theories about reason into current theories of cognition – or, even by failing in this attempt of combination; it may teach us where the two approaches have to part ways without an option of reconciliation.

I would like to highlight some critical aspects close to the idea of a logical faculty. Hanna justifies the logical faculty thesis by an extension of Chomsky's learnability arguments for natural languages and by a Kantian argument seeing in the logical faculty the transcendental condition for understanding any specific logical system. Both are strong arguments. Hanna *de facto*, although not explicitly so, engages in a on-going debate between philosophical logicians whether there is or can be one universal logic or whether there is irreducible logical pluralism. A recent statement of pluralism is Greg Restall's and JC Beall's *Logical Pluralism* (cf. [17]). Their critics – and Hanna may join in here – argue that even if there is a plurality of systems we are able to understand them all, and we are able to argue about them. These arguments have to use, it seems, some common logic. Hanna does not say much what belongs into the protological faculty. He mentions only some basic principles like the concept of validity, but also the highly controversial principle of non-contradiction. He believes that to identify more is not the task of philosophers like him. Getting to work on this task, however, may be the cardinal way to verify the idea of a universal logical faculty. Logical universalists have therefore begun to work out specific systems that can be used either as universal logic or as fall back system while using more than one system.[3] If there is such a universal logic (or something like Hanna's 'logic of thought') it can in at least one crucial aspect not be like universal grammar (or the *language of thought*). The prin-

ciples of universal grammar are cognitively inaccessible, at most some of the rules of individual natural language grammars are tacitly known and open to some limited cognitive access. The linguist comes up ex post with her theory by explaining the verbal behaviour and judgements on wellformedness by competent speakers (cf. [2]). This cannot be true with respect to principles of logic and rationality, since we not only follow these rules, we also represent them explicitly in processes of deliberation and argumentation to ourselves. Rules of logic are much more like semantic knowledge, which has to be cognitively accessible in verifying or rejecting statements. Hanna claims that we have a capacity of 'logical intuition', but that stays, despite Hanna's phenomenological claims on feelings of 'doxastic ease' and 'a sense of rational guidedness' in working with imagined symbols, as mysterious as similar claims by logicians like Gödel or Brouwer. One needs rather a theory of making our tacit logical knowledge explicit.

A similar criticism applies to Douglas Patterson's ([16]) theory of inconsistent semantics. Patterson's theory is a version of 'the inconsistency view' (that natural language semantics is inconsistent). He combines the idea that the paradoxes are derivable contradictions (an idea famous in some quarters of paraconsistent logic) with the linguistic idea of a lexical module within the language faculty: 'understanding a language can be a relation to a false semantic theory' (p. 198). Since believing falsehoods one knows to be false is irrational, the theory has problems of self-appraisal. How can anyone – Patterson included – believe a semantic theory just shown to be false? To circumvent this self-defeat Patterson has to endorse a modularity view of semantics: semantics is not just mainly implicit (sub-doxastically processed), but is a module in the strict Fodorian sense (cf. [11]). Our theories cannot cognitively penetrate the stored semantic knowledge: 'understanding a natural language is sub-doxastically cognizing a semantic theory that the paradoxes show to be logically false' (p.221). As, according to Patterson, semantics is cognitively impenetrable, we do not engage in irrationality.[4]

Nonetheless it seems unavoidable that we trust in sharing semantic knowledge with our interlocutors, semantic knowledge that we can appeal to if the occasion (of using "rectangular" or "quadratic", say) demands it.

Both theories neglect that speakers have to have some doxastic access to their internal semantic knowledge. A speaker has to have some knowledge of semantic rules, and this includes metarepresentation as one has to know whether in the light of conflicting information a term can

still be employed (cf. section 2).

5 Conclusion

Notwithstanding the ubiquity of metarepresentation in human cognition and the growing research on metacognition, many resulting questions close to procedural and computational models of cognition have to be more thoroughly explored. Precise models of the inner workings of our metarepresentational faculties are scarce. They require a general account of quotation, and an account of general (most non-conscious) cognition interacting with the language faculty and the lexicon. The aim of the present paper was to explore in parts the general role and some specific functions of meta-representations, and pose some representationalist questions about their formal mechanisms and syntax. An overall theory of a 'metamind' even if on the horizon still escapes our grasp.

Notes

1 I only look at the two reading where the whole clause is *de dicto* or *de re*. Of course there are mixed forms in which only some constituent is *de dicto* [see also section 4 below].

2 'Spell Out', 'Logical Form' and ¿Minimalism' are capitalized as they are meant to be understood as in the linguistic framework of 'Generative Grammar', for an introduction to their proper understanding cf. [13].

3 Prominent are several approaches rooted in the development of paraconsistent logics (cf. [5, pp.221-240]), especially Ross Brady's book *Universal Logic* (cf. [4]). There has even been a first world congress of universal logic in Montreux 2005, using a slightly different understanding of universal logic though, issuing in a new journal *Logica Universalis*.

4 But note that as long as the underlying logic is not paraconsistent we not only have an inconsistent sub-doxastic semantics, we have a *trivial* semantics endorsing anything by *ex contradictione quodlibet*.

Manuel Bremer
Institute of Philosophy
University of Düsseldorf
Universitätsstraße 1
40225 Düsseldorf, Germany

<bremer@mbph.de>

References

[1] Mark Baker. *The Atoms of Language. The Mind's Hidden Rules of Grammar*. Basic Books, New York, 2001.

[2] Cedric Boeckx. *Linguistic Minimalism. Origins, Concepts, Methods, and Aims*. Oxford University Press, Oxford, 2006.

[3] Laurence BonJour. *The Structure of Empirical Knowledge*. Harvard University Press, Cambridge, MA, 1985.

[4] Ross Brady. *Universal Logic*. CSLI Publications, Stanford, 2006.

[5] Manuel Bremer. *An Introduction to Paraconsistent Logics*. Peter Lang, Bern, 2005.

[6] Herman Cappelen and Ernie LePore. *Language Turned On Itself. The Semantics and Pragmatics of Metalinguistic Discourse*. Oxford University Press, Oxford, 2007.

[7] Donald Davidson. Deception and division. In J. Elster, editor, *The Multiple Self*, pages pp.79–92. Cambridge University Press, Cambridge, 1980.

[8] Donald Davidson. *Inquiries into Truth and Interpretation*. Oxford University Press, Oxford, 1984.

[9] John Dunlosky and Janet Metcalfe. *Metacognition*. CA, Thousand Oaks, 2009.

[10] Robert Fiengo and Robert May. *De Lingua Belief*. The MIT Press, Cambridge, Mass., 2006.

[11] Jerry Fodor. *The Modularity of Mind*. The MIT Press, Cambridge, Mass., 1983.

[12] Robert Hanna. *Rationality and Logic*. The MIT Press, Cambridge, Mass., 2006.

[13] Norbert Hornstein, Jairo Nunes, and Kleanthes Grohmann. *Understanding Minimalism*. Cambridge University Press, Cambridge, 2006.

[14] David Kaplan. Quantifying in. In Donald Davidson and Jaakko Hintikka, editors, *Words and Objections: Essays on the Work of W.V. Quine*, pages pp.178–214. Reidel, Dordrecht, 1969.

[15] Keith Lehrer. *Metamind*. Oxford University Press, Oxford, 1990.

[16] Douglas Patterson. Understanding the liar oxford. In JC Beall, editor, *Revenge of the Liar. New Essays on the Paradox*, pages pp.197–224. Oxford University Press, Oxford, 2007.

[17] Greg Restall and JC Beall. *Logical Pluralism*. Oxford University Press, Oxford, 2006.

[18] Dan Sperber, editor. *Metarepresentations. A multidisciplinary perspective*. Oxford University Press, Oxford, 2000.