Scope of Semantic Innocence

Jaya Ray
Assistant Professor
Lakshmibai College, University Of Delhi
India
An Introduction to
ATINER's Conference Paper Series

ATINER started to publish this conference papers series in 2012. It includes only the papers submitted for publication after they were presented at one of the conferences organized by our Institute every year. The papers published in the series have not been refereed and are published as they were submitted by the author. The series serves two purposes. First, we want to disseminate the information as fast as possible. Second, by doing so, the authors can receive comments useful to revise their papers before they are considered for publication in one of ATINER's books, following our standard procedures of a blind review.

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research
This paper should be cited as follows:

Scope of Semantic Innocence

Jaya Ray
Assistant Professor
Lakshmibai College, University Of Delhi
India

Abstract

The principle of semantic innocence demands that a particular linguistic expression should have the same semantic value irrespective of the linguistic contexts in which it appears. But Frege, by postulating a two-tier theory of meaning for the same word, seems to have violated semantic innocence. But why did Frege do so? An attempt to answer this question reflects that propositional attitude ascribing contexts, specially belief contexts, pose a serious threat to the principle of semantic innocence by breaking another very important principle of language – the principle of substitutivity. According to this principle, two co-referential terms are substitutable salva varitate. But the principle of substitutivity seems to disappear in belief contexts. In belief contexts, if a term is substituted by another co-referential term, the truth value of the sentence might change, which shows that the two co-referential terms make different semantic contributions to the sentences in which they occur and that is a clear violation of the principle of semantic innocence.

To face this challenge I have tried to analyse the logical form of belief sentences, since semantic interpretations are largely attached to the level of logical form. My analysis of belief sentences reveals that the problem is not really a problem with de-re beliefs, but only with de-dicto beliefs. A two-tier Fregean conception of semantics violating the condition of innocence may definitely apply, if at all, in propositional attitudes concerning belief ascribing sentences in de-dicto (logical) form. The rest of the language seems to be covered by referential semantic innocence.

Key Words:

Corresponding Author:
The principle of semantic innocence demands that a particular linguistic expression should have the same semantic value irrespective of the linguistic contexts in which it appears. The importance of this principle lies on the fact that it is also a requirement of a child’s language acquisition capacity. This requirement arises from an adequacy condition posed on a putative theory of language. The condition of explanatory adequacy demands that a theory of language must explain how is it possible for virtually all human beings, regardless of intelligence, motivation or even desire, to acquire the native language so regularly and easily.

A theory of language primarily consists of a theory of grammar and a theory of meaning. It is already an established fact by now that a theory of grammar is universal, uniform and single, in other words, it is innocent; otherwise it is not possible for a two-or-three year old child to acquire the complex structure of her native language. So the requirement of explanatory adequacy needs to be carefully addressed in the domain of a theory of meaning. How is it possible for a normal human child to figure out the meanings of as many as thousands of words from the complex noisy linguistic environment to achieve his estimated vocabulary? An attempt to answer this question reflects the need for an innocent semantics.

Frege’s Two-tier Theory of Meaning

Frege’s theory of meaning was first outlined in his article ‘Funktion und Begriff’ (1891) in regard to mathematical expressions, and was expanded and explained in greater details in his most famous work ‘Über Sinn und Bedeutung’ (‘On Sense and Reference’, 1892). In this latter article, Frege differentiated between the nominatum or reference [the designated object] of a sign [name, word combination and expression] and the sense [connotation, meaning] of the sign in which is contained the manner and context of presentation. For example, the nominata of ‘evening star’ and ‘morning star’ are the same, but not their senses.

Frege raised several problems that forced him to go beyond a theory of meaning in terms of reference only. His main intention of introducing the notion of sense of an expression was to address the problem of cognitive significance that arises in the context of

(1) identity statements,
(2) vacuous terms and
(3) propositional attitude ascriptions.

For the purpose of this paper I will deal only with Frege’s problem of propositional attitude ascriptions. This problem can be taken as a generalisation of Frege’s problem concerning indirect speech to a whole class of expression that is non-truth-functional. It is the puzzle concerning the apparent failure of substitutivity of co-referential singular terms in certain contexts of indirect
speech, particularly in propositional attitude ascribing contexts – contexts where the principle of substitutivity of co-referential terms fail.

**The Problem of Substitution**

The principle of substitutivity states that if two terms \( p \) and \( q \) are extensionally equivalent, then we can replace some or all occurrences of \( p \) by \( q \) in a statement \( F(\alpha) \) without changing its truth-value. Thus,

\[
p=q \Rightarrow F(p)=F(q)
\]

Normally this principle does not pose any problem. From the truth of

1. Hesperus = Phosphorus,
2. Hesperus is a planet,

and substitution, we can validly infer

3. Phosphorus is a planet.

However, there seem to be some serious counter examples to this principle. The principle of substitutivity holds only for those propositions which are truth-functional. But for non-truth-functional propositions like indirect speech the case is completely different. For example, we know that the names ‘Hesperus’ and ‘Phosphorus’ have the same reference, i.e., both of them refer to the planet Venus. Yet, it is not always true that they can replace one another without changing the truth of the proposition. For example, consider the following inference:

5. Riya believes that Hesperus is a planet.
6. Therefore, Riya believes that Phosphorus is a planet.

If we assume that Riya is unaware of the fact that Hesperus is the same planet as Phosphorus, then (6) becomes false, though (4) and (5) are true. Therefore, in this context we cannot substitute the name ‘Hesperus’ for ‘Phosphorus’ and arrive at a valid conclusion. One of the most important challenges in the philosophy of language is to give a satisfactory account of these contexts where the principle of substitutivity fails (these are known as opaque contexts). If the two co-referential names ‘Hesperus’ and ‘Phosphorus’ make different semantic contributions to the sentences in which they occur, which is a clear violation of the principle of semantic innocence, how will the child understand which theory of meaning to be used to interpret the meanings of those two sentences?
De re and De dicto beliefs

Traditionally speaking a de re belief is a belief about a particular individual or object that has certain property, whereas a de dicto belief is a belief in a certain proposition (of what is stated). For example, I can believe of the earth that it moves round the Sun; and I can also believe that the earth moves round the Sun. In the first case it is a belief de re and in the second case it is a belief de dicto. In the first example, the belief ascription is made by a relational statement involving myself, the earth and its movement round the Sun, and therefore, no proposition seems to be involved here; but in the second one it is clearly a belief in a proposition, being specified by the ‘that’ clause, as noted.

It can be said that every belief is de dicto in a basic sense, i.e., in the sense that it has a propositional content, but it is purely de dicto if and only if an accurate expression of its propositional content is a quantified statement in a sense to be specified below; otherwise it is de re. That is why, there is a special way of reporting a de re belief by saying explicitly that the belief is of an object, such as I believe, of the earth, that it moves round the Sun.

Thus Quine pointed out the ambiguity in the following sentence: (7) Ralph believes that someone is a spy.

This could mean either of the following:

(8) Ralph believes that there are spies.
(9) There is someone whom Ralph believes to be a spy.

The first is the de dicto sense of believing in which there is a dyadic relation between a believer and a proposition. While in the second case, there is a de re sense of believing in which there is triadic relation among a believer, an object, and a property. The distinction between the two can also be seen as a distinction of scope for the existential quantifier. In (8), the existential quantifier is interpreted as having narrow scope, within the scope of ‘believes’:

(10) Ralph believes that (∃x) (x is a spy).

In (9), the existential quantifier has wide scope and binds a variable that occurs freely within the scope of ‘believes’:

(11) (∃x) (Ralph believes that x is a spy).

Accordingly we can say that the de re / de dicto distinction can be analysed from a syntactic and a semantic point of view. A sentence can be said to be syntactically de re if it contains a free variable within the scope of an opacity verb that is bound by a quantifier outside the scope of that verb. Otherwise it is syntactically de dicto. So, (10) is syntactically de dicto, whereas (11) is syntactically de re.
Again a sentence is said to be semantically \textit{de re} if it permits substitution \textit{salva veritate}, which means, if the terms or phrases occurring within the sentence can be freely substituted by co-referential terms without altering the truth-value of the sentence. Otherwise it is said to be semantically \textit{de dicto}. Thus, \textit{de re} belief attributions are referentially transparent whereas \textit{de dicto} belief attributions are referentially opaque. For example, if I believe \textit{de re} of Hesperus that it is a planet,

then we can substitute any co-referential term for Hesperus, regardless of whether I would describe Hesperus in that way. That is why \textit{de re} believes are referentially transparent. But if I believe \textit{de dicto} that Hesperus is a planet, then the substitution of Hesperus by any co-referential term fails.

Ever since Frege raised the problem of substitution in opaque contexts, philosophers of language and linguists, especially formal semanticists, have been trying to address the problem by specifying the logical form of these contexts. According to Russell, what constitutes the puzzle about the nature of belief is that quite often we have false beliefs. He took the example of a false belief like ‘Othello believes that Desdemona loves Cassio’, where the subordinate verb ‘loves’ seems to occur as relating Desdemona to Cassio, but in fact it does not do so. That is why Russell holds, ‘You can not get in space any occurrence which is logically of the same

form as belief’ (Russell, 1918, p. 82). And by ‘logically of the same form’ he means that ‘one can be obtained from the other by replacing the constituents of the one by the other ’. Russell thus considers beliefs as new type of propositions – ‘new beast for our zoo’– different from every other proposition.

It is obvious that the postulation of ‘new beasts for the zoo’ of facts raises exactly the same problems for explanatory adequacy as with non-innocent
tories. If belief-facts are not to be understood with facts that are presented in terms of acquaintance (Russell’s phrase), how does the child know that in interpreting a ‘belief’-sentence he now is confronted with a logically ‘new fact’ for which he needs to marshal wholly new sort of evidence? It seems that appeal to ‘new facts’ simply labels Frege’s problem without addressing it.

Philosophers of language and formal semanticists like Davidson (1968),

Higginbotham (1986, 1991), Larson and Ludlow (1993), Barwise and
Perry (1987) and many others have addressed the issue which has generated a vast literature covering a variety of formal approaches with rich resources; but that is beyond the scope of this paper. All I can say is that these theories either re-state the problem, or address one side of the problem for \textit{de dicto} beliefs. The issue of whether \textit{de dicto} beliefs can be given satisfactory logical forms at all still remains unanswered.

Therefore, we have the choice left between accepting either the sceptical
solution that there can not be a theory of meaning for belief sentences [like Kripke (1979), Schiffer (1987)] or looking at a new horizon for the analysis of belief sentences so that we can analyse why it is not possible to get a logical structure of belief sentences. This paper chooses the second option because if there is no theory of meaning for belief sentences then we owe an explanation to the fact that we do understand the meanings of those sentences involving
beliefs. So something must have gone wrong with our process of analysis. The only justice we can do with the theory of language as well as belief sentences is to change our attitude towards beliefs by treating belief as a ‘local’ concept, in the sense explained in this work.

**Belief as a Local Concept**

Akeel Bilgrami, following Davidson, draws a close connection between meaning and belief. For him, the literal meaning of a sincere, non-self-deceived, utterance of a sentence by an agent gives the content of the belief that is expressed by that utterance. The central thesis of Bilgrami’s philosophy is a very specific *externalist* picture of content comprising of two sub-theses – *the unity of content and the locality of content*.

Bilgrami’s proposed externalism consists of a general notion of the doctrine of externalism – that the contents of an agent’s beliefs are not independent of the world which is external to the agent, i.e., intentionality can not be completely characterized independent of the external world – added with a constraint that gives a specific reading to the general doctrine of externalism. The constraint is stated as: ‘When fixing an externally determined concept of an agent, one must do so by looking to indexically formulated utterances of the agent which express indexical contents containing that concept and then picking that external determinant for the concept which is in consonance with other contents that have been fixed for the agent’ (Bilgrami, 1991, p 5).

Several things should be made clear in this constraint thesis before going into the explanation of Bilgrami’s theses the *unity of content* and the *locality of content*. First of all, Bilgrami points out that ‘concept’ is here used as the counterpart to ‘term’ in the same way as ‘content’ is thought of as the counterpart to ‘sentence’. Secondly, as Bilgrami takes concepts to be externally constituted and contents are composed of concepts, it follows that the contents of an agent’s mind are necessarily public. This amounts to saying that meanings and belief-contents are public phenomena in the sense that they are properties of the agent that is available, in principle, to other agents, who have the capacity for meaning and belief, who can experience the same external environment that constitutes, in general, an agent’s meanings and beliefs. The publicness of belief and meaning is further explicit from the fact that fixing of an agent’s concepts and contents are done by others. Thirdly, the fixing of externally determined concepts takes place, by looking at the indexical contents in which they occur and the utterances expressing them. Indexical contents and utterances, therefore, provide us with essential clues to an agent’s perceptions and responses to things and show us the path to reach others’ mind.

Bilgrami’s thesis of the *unity of content* claims that there is only one notion of content that is externally determined and that very notion explains actions. It rejects the orthodox distinction between the ‘wide’ and ‘narrow’ content according to which if content is externally determined (‘wide’ content),
then it can not go into the explanation of commonsense psychological behavior, and for that another internal notion of content (‘narrow’ content) is needed. But Bilgrami’s thesis is a unified notion of content as it states that though contents are externally determined, it is the same contents that can be used in the commonsense psychological explanation or rationalization of behavior. The thesis of unity of content basically tries to put together two widely held assumptions regarding intentional content. The first assumption is that intentional states have a role in the commonsense psychological explanation of action and behavior, and the second assumption is that these contents are externally determined.

While analyzing the reason why philosophers have opposed his thesis of unity of content, Bilgrami finds a hidden tension between the above two assumptions. The tension is due to the fact that since the purpose of the explanation of an agent’s action and behavior is to illuminate the behavior of the agent itself, whatever is cited in an explanation must be constituted within the agent and therefore, the examples should show the explanatory failure of externally constituted contents. That is why it is argued that two notions of content are needed. Bilgrami has analyzed the following example to show why he is in favour of a unified theory of content rather than a bifurcation of content.

The example he cited is Kripke’s well-known puzzle about belief, but rather than taking up Kripke’s original puzzle, Bilgrami here works with a variant of Kripke’s example suggested by Brian Loar (1985). In this variant a person named Pierre learns from his nanny, while growing up monolingually in Paris, something which he expresses by saying ‘Londres est jolie’. Later on, he settles in a pretty section of London and after picking up the native language he is disposed to say, ‘London is pretty’. He, however, does not realize that the city he learnt about in Paris is the same city he is living now.

From the close connection between the meaning and belief mentioned at the beginning of this section we can infer that if Pierre is disposed to utter or assent to ‘London is pretty’, then he believes that London is pretty. Kripke formulates this assumption through his principle of disquotation. The question that occurs at this point is: does Pierre’s belief have the same content when he assents to ‘Londres est jolie’? According to Bilgrami, those, who will answer this question in the affirmative to show the explanatory failure arising out of externally determined content, will use two further assumptions apart from that of the principle of disquotation. One is that externalism is tied to a certain notion of reference, and the other is that translation preserves reference and therefore the truth of the sentence that is assented to. For them, since ‘London’ refers to London and since ‘London’ translates ‘Londres’, the two belief contents must be the same.

Kripke’s original example gives rise to a puzzle by attributing inconsistent beliefs to Pierre (Kripke 1975). Bilgrami shows that Pierre has the same belief attributed to him earlier in Paris and later in London. The difference is due to the contingent circumstances where Pierre settles in London. But the source of an element involved in both the examples, which is independent of the
contingencies of the examples and which has nothing to do with Pierre being attributed inconsistent beliefs, is that Pierre does not know about the identity that holds between ‘London’ and ‘Londres’. This common element produces a more general difficulty: ‘Belief attributed, under the assumptions mentioned, simply cannot account for the inferences that an agent such as Pierre might make or fail to make … these examples bring out the fact that content attributed under these assumptions will not always be efficacious in explanations’ (Bilgrami, 1992, p. 17).

According to Bilgrami, these content attributions are ineffectual due to the insensitivity to Pierre’s different conceptions of London. Pierre conceives of London differently when he is in Paris and when he is in London. And therefore, it is possible for him to fail to know or realize the identity of London and Londres.

Bilgrami’s solution to this problem lies in his proposal of the constraint mentioned earlier which he adds to his theory of externalism. The constraint demands that the external items correlated with an agent’s concepts be described in such a way that it is in consonance with the other contents attributed to the agent and have the effect of including his conceptions of the external items. This constraint makes it sure that we get the right description of the external items to determine an agent’s concepts. For example, if we allow the external elements to enter content under this constraint, the Pierre example will not have any bifurcatory consequences. In Pierre’s case, the city in question, which is correlated with Pierre’s earlier use of the term ‘Londres’ and later use of the term ‘London’ will be under different beliefs or descriptions of his. In Bilgrami’s notion of externalism, one looks at the agent’s relations with the world around him with the task of imposing characterizations in which the concepts that we find composing the agent’s contents can also be used to describe the world around him with which he relates.

Bilgrami’s thesis of the locality of content is a perfect supplement to his idea of the unity of content. By locality of content, Bilgrami means to say that the explanation of an agent’s actions always takes place at a much more local level than the meaning-theoretic level. The meaning-theoretic level is the level where the theories of meaning do their work. For example, theory of meaning for the term ‘water’ in a particular agent’s language can be seen as an ideally complete specification of his ‘concept’ attributed to him on the basis of all the beliefs, which he associates with ‘water’. But the entire aggregate of beliefs is not required for the attribution of specific content to explain behavior and here comes the importance of the locality of content. At the local level one sifts out of the pool of resources consisting of various beliefs only those which are relevant in order to attribute specific contents for the explanation of action.

Thus, it is very important to understand that a distinction needs to be made between two levels – the level of a theory of meaning, on the one hand, where all the beliefs are aggregated in the specification of a given concept; and the level of content, on the other hand, where actions are explained. In the explanation of an action, belief-contents are invoked, and these belief-contents are composed of concepts, but *not of all the concepts specified at the meaning-
theoretic level, rather only of those concepts selected at the local level.

For our purposes, Bilgrami’s handling of the problem of belief attribution leads to several crucial consequences.

A. Bilgrami’s position with respect to content of beliefs, namely, unity of content, is a species of semantic innocence advocated throughout this work. Bilgrami is opposed to the distinction between narrow and wide contents, which, in semantic terms, translates into a distinction between Fregean (narrow) and Russellian (wide) semantics (Fodor 1994). Bilgrami not only rejects the distinction, he holds that the content of belief is basically ‘externalistic’ in character. To that extent he is not only advocating semantic innocence, he is opting for referentialism as the only candidate for semantic innocence.

B. Bilgrami advocates a central and novel distinction between ‘meaning-theoretic’ explanation and explanation of content, especially for contents of beliefs. This is where Bilgrami’s suggestions take a decisive sceptical turn because, in effect, he is arguing that the contents of beliefs escape meaning-theoretic explanation. In so far as theories of logical form are viewed as meaning-theoretic explanations, contents of beliefs cannot be captured in terms of the form of belief-propositions. Translated in our terms, semantic innocence in belief-contexts is not compatible with their logical form.

Concluding Remarks

What then is the locality of content in belief-ascriptions? According to Bilgrami, ‘the entire aggregate of beliefs is not required for the attribution of specific content to explain behavior and here comes the importance of the locality of content’. Bilgrami’s solution requires that something like the notion of ‘what Pierre has in mind at t’ determines the content of Pierre’s beliefs at t, where ‘has in mind at t’ is to be understood locally. That is, the primitive notion for determining the content of Pierre’s beliefs is not the semantic value of one of its constituents (for example

‘London’), however constrained at the circumstance of evaluation, but that content of ‘London’ that Pierre has at t. Once the effect of locality is understood in this way, the account of Pierre’s beliefs becomes a ‘psychological account’ in Schiffer’s sense, although we can continue to assume that each of these states of Pierre are driven by externalistic content. In other words, there is no necessary shift to ‘narrow content’ just because the notion of what Pierre has in mind has been taken to be the primitive. However, as Schiffer’s incisive analysis revealed, any psychological account of the sort just described is incompatible with the demands of a (general) theory of meaning that is supposed to hold for a language, not just for Pierre’s current beliefs. Therefore, given the link between the locality conditions governing Pierre’s psychological states, and the limits of formal enquiry on the content of those states, we can see why a theory of beliefs does not fall under putative theories of meaning for language that we can currently make sense of.
Generalizing to Frege’s example, we can now say that in evaluating S’s belief that Hesperus is a morning star, we need not simultaneously take into account S’s other belief that Phosphorus is an evening star. In each case we need to treat S’s use of ‘Hesperus’ and ‘Phosphorus’ separately in terms of the locality conditions governing their use, even if in meaning-theoretic terms Hesperus = Phosphorus. Therefore, S cannot be viewed as saying, of the same star, that it is at once a morning and an evening star; the locality conditions governing S’s use of these names does not guarantee that one can be substituted for the other. By parity of reason, it cannot also be ruled out that S might be so governed by locality conditions that he holds that Hesperus = Phosphorus (just as we do!). In that case S is not likely to say of the same star that it is at once a morning and an evening star. If, despite these precautions, S does in fact say so, he can legitimately be charged for holding (apparently) inconsistent beliefs. But these decisions depend entirely on S’s mental biography, rather than on the meanings of terms.

We could say that we might have entertained a philosophical view that contents of beliefs also fall under general theories of language in terms of valuation of the predicate ‘believes that’ because giving a theory of meaning for an expression X is the only way of explaining what we understand on saying/hearing X. The fact that the explanatory scope of our favourite theories of meaning – referentialism, in this case – can be extended to cover embedded ‘that’-clauses – such as ‘sees that’ and ‘says that’ – might have encouraged us to extend the scope further to ‘believes that’ as well. After all, if language is used to talk about the world, then the general conditions governing that talk (=theories of meaning) ought to cover every piece of talk including those involving ‘believes that’; how else do we make sense of that talk?

The locality of content idea articulated above can be used against the generality of content implicit in the preferred philosophical concept of belief. If this approach is plausible, then the problem with de dicto belief proposition and the substitution problem thereof could be traced to the generality assumption implicit in the philosophical concept of belief (Mukherji, 2006). We can thus say that the problem may not arise if we give up the generality requirement and with it the meaning-theoretic requirement – that is the requirement of logical form – imposed on belief de dicto.

The prevalence of the locality condition in determining the content of belief reports opens new directions for philosophical clarification of the concept of belief. Although beliefs are reported in language, the content of belief do not seem to obey the classical conception of semantics as language-world connections, where the ‘world’ is typically understood as the external, mind-independent world – referentialism. The locality condition suggests that what rules in belief contexts is the ‘constructed world’ of the agent of beliefs – the ‘thought’ world. This was Frege’s central insight. But the (mind-internal) world that essentially determines the content of beliefs does not pervade the entire length and breadth of language – the locality; hence, it defies the classical conception of semantics. It could be that the ‘world’ of beliefs is a different kind of world; it is my world, as Wittgenstein would say
(Wittgenstein, 1961; Boruah, 2006). A study of this insight is beyond the scope of the present project.

References


