A Case for Reference to Phenomenal Experience in Natural Language

Gaetano Fiorin\(^1\,2\) & Denis Delfitto\(^2\)

\(^1\)Utrecht Institute of Linguistics, Utrecht University
\(^2\)Department of Philology, Literature, and Linguistics, University of Verona

\(<g.fiorin@uu.nl, denis.delfitto@univr.it>\)

Resumé

Model-theoretic semantics is grounded on the assumption that meaning is reference. In such a framework, defining the meaning of a natural language expression corresponds to defining the object to which it refers, a practice sometimes referred to as ‘Natural Language Metaphysics’. Typical objects of reference used in semantic modeling are individual entities and possible worlds. A number of linguistic facts have been explained by introducing novel classes of entities in the model. Relevant examples are times and events. In this contribution, we aim at producing a case for introducing phenomenal experiences among the set of objects natural language expressions refer to. We show that the truth conditions of some natural language expressions are best described as denoting phenomenal experiences, rather than objects of the world.

Mots clés: phenomenal consciousness, implicit de se, immunity to error through misidentification, private language, evidentiality

1. Model-theoretic truth-conditional semantics

Richard Montague (Montague 1970, 1973) set the grounds for the development of formal semantics by applying to natural language the tools originally developed in philosophy, logic and mathematics for the analysis of formal languages. The theory of linguistic meaning developed by Montague has two main components. The first is a set of rules for translating natural language expressions into formulas of intensional logic, expressing truth-conditions. The second is a model theory for the evaluation of the truth of formulas within a model. The first component has been at the basis of much interdisciplinary work at the interface between natural language semantics and syntax. The second component, the model theory, is primarily aimed at explaining

\(^1\) We wish to thank the organizers and participants of the LogPrag Workshop (17-19 juin 2015, Les Diablerets, Switzerland) for their many useful comments and suggestions. We are also very grateful to Anne Reboul and Floris van der Burg for discussing with us a number of issues presented in this article. The usual disclaimers apply.
the notions of truth and entailment. The truth of formulas of intensional logic is evaluated in a model, typically conceived as a set of objects, also called a domain, and an interpretation function, assigning to each expression in the language of intensional logic an object in the model. A sentence is true if and only if (iff) the truth conditions expressed by the sentence are met by corresponding set theoretic relations between the objects in the model. The study of this second component has been at the basis of a profitable exchange between semantics and philosophy, in particular in defining the nature of the objects that should be included in the domain of the model. The model originally developed by Montague includes two main classes of objects: individual entities and possible worlds. Montague regards propositions as sets of possible worlds and, accordingly, translates the notion of entailment into the set-theoretic notion of inclusion of one set of possible worlds into the other.

Montague’s framework is grounded on the assumption that meaning is reference. According to this view, the meaning of a natural language expression corresponds to the object to which it refers. This view has become predominant in semantic theorizing particularly after the work of Saul Kripke and David Kaplan, who provided decisive arguments against the Fregean analysis of proper names and indexicals (Kaplan 1989, Kripke 1980). Meaning is here defined as a relation between an expression of the language and an object of the world. It should be observed that the externalist nature of the hypothesis, whereby meaning is a mapping between expressions of a language and objects of the world, inevitably minimizes the subjective role of the interpreter in the construction of linguistic meaning: "[…] reference […] depend[s] on relations between the speaker and his social and physical environments that are best understood not by investigation the speaker’s mental repertoire but by inquiring into the chain of circumstances that led the speaker’s acquisition or present use of the name” (Burge 1992, 24).

2. Natural language Metaphysics

According to the model described in the previous section, a significant part of the semanticist work consists in providing a model of the world, conceived as a set of objects that function as the reference of linguistic expressions. Bach (1986) has labeled this practice ‘natural language metaphysics’, a term chosen to carefully distinguish it from the metaphysics canonically practiced in philosophy: Whereas metaphysics studies how things are, natural language metaphysics is concerned with what kinds of things people talk as if there are. As mentioned above, the model originally devised by Montague includes two classes of objects: individual entities and possible worlds. In
Bach’s terms, Montague’s proposal amounts to the hypothesis that people talk as if the world was organized in individual entities and possible worlds.

Part of the task of the semanticist has consisted, and still consists, in individuating the model, and the corresponding natural language metaphysics, that provides the best account of linguistic meaning. Whereas it is common practice to maintain the number of object classes in the model as small as possible, a number of arguments have been produced in the literature in favor of the introduction of novel classes. To mention but a few, Montague himself was already adopting reference to points in time to account for tense in English, Davidson (1967) and Parsons (1990) provide arguments for adopting events as a distinct class of objects, Barwise (1981) and Barwise and Perry (1983) support the introduction of situations as yet another, independent class. In this contribution, we will produce an argument that a further class of objects should be introduced in the model: phenomenal experiences.

3. Attitudes de re and de se

Hintikka (1962) offers a framework for the representation of beliefs based on the notion of doxastic alternative. Let $a$ be an intentional agent and $w$ a possible world. Let $Bel(a,w)$ be the set of possible worlds $w'$ such that $a$, on the basis of the knowledge she has in $w$, has no reason to exclude $w'$ as a candidate for being the actual world. Let us call $Bel(a,w)$ the set of doxastic alternatives that $a$ has in $w$. We say that $a$ believes a proposition $p$ iff $a$’s doxastic alternatives entail $p$, that is, iff the set of possible worlds corresponding to $Bel(a,w)$ is a subset of the set of possible worlds corresponding to $p$. Accordingly, a belief report such as (1) is taken to denote a relation between a subject, Jim, and a proposition, $p$, which is satisfied in $w$ iff Jim’s doxastic alternatives in $w$ entail the proposition $p$:

(1) Jim believes that $p$
(2) $Bel(j,w) \subseteq p$

Hintikka’s framework can be extended to cover other propositional attitudes in the same fashion, that is, as relations between two sets of possible world: a modal base defined relatively to the subject of the propositional attitude and a proposition corresponding to the object of the propositional attitude.

As is well know, however, the framework fails to provide a satisfactory account of two types of propositional attitudes: attitudes de re and attitudes de se. A de re report is a report of an attitude that an agent has about a certain individual or object, that is, a res:
A subject can also have an attitude *de re* towards an individual that happen to be the subject herself. To see an example, sentence (4) can report, in one of the available interpretations of the pronoun *she*, an attitude that Pamela has towards herself. Furthermore, we can distinguish between the case in which Pamela has an attitude toward herself without being aware of the identity between the subject and the object of her attitude and the case in which Pamela has an attitude toward herself being fully aware of the identity between herself and the object of her attitude. We will refer to attitudes of the former type as attitudes *de se* (Lewis 1979). Importantly, natural languages display several means for unambiguously reporting a *de se* attitude or speech. Compare (4) with (5).

Whereas (4) is compatible both with a scenario where Pamela is aware of the identity between herself and the object of her belief and with a scenario in which she is not aware of such identity, (5) is only compatible with a scenario in which Pamela is aware of the identity between the *res* of her belief and herself.

As anticipated, Hintikka’s semantics of belief reports does not offer a satisfactory account of *de re* and *de se* beliefs. The reason is, in short, that in Hintikka’s framework propositional attitudes are relations between a subject and a propositions, whereas attitudes *de re* and *de se* seem to correspond to relations between a subject (the holder of the propositional attitude), a *res* (the object of the propositional attitude) and a property that is attributed to the *res*. To see an example, consider again sentence (5). Hintikka’s account assigns to (5) the truth conditions in (6), which are verified iff Pamela’s doxastic alternatives entail the proposition that Pamela looks beautiful.

The truth-conditions in (6), however, are verified independently of whether Pamela is aware of the identity between herself and the object of her belief.

Cresswell & von Stechow (1982) offer a comprehensive solution for the problem of attitudes *de re* and *de se* based on two main ingredients. The first is *centered possible worlds*. Following Lewis (1979), they propose that the object of propositional attitudes is not a set of possible worlds (that is, a proposition), but a set of (ordered) pairs of possible worlds and individuals, \(<w, a>\), that is, a set of *centered possible worlds* (the terminology is from Quine 1969). The intuitive idea is that the object of a propositional attitude is not just a possible world...
but a possible world as seen through the point of view of an agent, that is a center. Technically, the notion of doxastic alternative and the truth conditions of belief reports can be redefined on the basis of the notion of centered possible world. Firstly, a centered possible world \( <w, b> \) is a doxastic alternative for \( a \) in \( w \) if \( a \)'s belief state does not rule out the possibility that \( a \) is the individual \( b \) living in world \( w \). Let us refer to \( a \)'s doxastic alternatives in \( w \) as Lewis-doxastic alternatives, or \( \text{Bel}_L(a, w) \). Secondly, \( a \) believes that \( \omega \), where \( \omega \) is a set of centered possible worlds, iff \( a \)'s doxastic alternatives entail \( \omega \), that is, iff \( \text{Bel}_L(a, w) \subseteq \omega \).

The second ingredient in Cresswell & von Stechow’s proposal is Kaplan’s (1969) notion of acquaintance relation. More precisely, Cresswell and von Stechow propose that an agent \( a \) believes de re of \( b \) that \( P \) iff the following three conditions are met:

1. \( R \) is a contextually relevant acquaintance relation;
2. \( R \) relates the agent of the belief \( a \) with the res \( b \);
3. \( \text{Bel}_L(a, w) \subseteq \{<w, x>: y \text{ is the object } x \text{ is related to by } R \text{ and } P(y) \text{ is true in } w\} \).

In this framework a de re belief is a belief that an agent has toward a certain res that is individuated on the basis of a certain acquaintance relation that the agent has towards the res. A de se belief is a special kind of de re belief where the acquaintance relation is one of self-identification. When Pamela has a de se belief about herself, she believes that the individual she is acquainted with as being identical to herself has the property of looking beautiful.

In conclusion, centered possible worlds and acquaintance relations can successfully account for de se belief reports. Importantly, they allow us to do so without having to introduce a novel class of objects in our model. After all, centered possible worlds are relations between individuals and possible worlds and acquaintance relations are relations between individuals.

In the next section, we will present a type of de se, originally discussed in Higginbotham (2003), that cannot be reduced to an account in terms of centered-possible worlds and acquaintance relations.

4. Remembering and imagining

Higginbotham (2003) observes that the silent subject of the infinitival complements of the verbs remember and imagine is (i) unambiguously de se and (ii) immune to error through misidentification relatively to the subject of the matrix clause.

Consider the minimal pair in (10) and (11). They are both unambiguously de se in that they report a memory that Ralph has
about himself, in full awareness of the identity between himself and
the object of his memory.

(10) Ralph remembers going to the movies
(11) Ralph remembers himself going to the movies

There is, nevertheless, a sense in which (10) is more “first-
personal” than (11). Firstly, as discussed by Fodor (1975), (14) follows
naturally from (12) and (13), whereas (15) does not. The fact that only
Churchill gave the speech and that he remembers doing so, does not
support the conclusion that only Churchill remembers himself giving
the speech. It could in fact be the case that someone else remembers
Churchill giving the speech. However, the same premises support the
conclusion (15).

(12) Only Churchill gave the speech
(13) Churchill remembers giving the speech
(14) Only Churchill remembers giving the speech
(15) Only Churchill remembers himself giving the speech

Secondly, the gerundive complement of remember contrasts with its
finite complement in that it has an event-like interpretation, rather
than a proposition-like interpretation. Suppose that my grandfather
died before I was born. In this scenario, I can remember that my
grandfather was called “Rufus” but I cannot remember my
grandfather being called “Rufus”. That is, (17) may be true in the
given scenario, but (16) is inevitably false.

(16) I remember my grandfather being called “Rufus”
(17) I remember that my grandfather was called “Rufus”

Finally, whereas (18) is perfectly acceptable, (19) is contradictory in
a way that is reminiscent of the “Moore Paradox”.

(18) I used to remember walking to school in the 5th grade, but I no longer
remember it.
(19) I used to remember that I walked to school in the 5th grade, but I no
longer remember it.

Intuitively, the crucial property of the control complement of
remember is that it reports the remembered event “from the inside”, as
suggested by Pryor (1999), that is, from the introspective, first-
personal point of view of the bearer of the memory. When I remember
going to the movies or walking to school, I remember the experience of
going to the movies or walking to school, and not just the fact that I
did so.

Similar observations can be made for the verb imagine:

(20) I imagined flying through space
(21) I imagined myself flying through space
Both sentences are de se, in that the subject must be aware that the object of his imagination is he himself. However, (21) does not specify whether I am imagining an object, which I identify with myself, flying through space or whether I am imagining the subjective experience of flying. The complement of (20), on the other hand, unambiguously refers to the subjective experience of flying through space.

For reasons of space, we restrict our presentation to the verbs remember and imagine. There are, however, reasons to think that this semantic property can be generalized to other control structures, as is the case, for example, for the infinitival complement of want (see Delfitto & Fiorin, in progress).

5. Immunity to error through misidentification

Higginbotham (2003) argues that the central feature of the gerundive complements of remember and imagine is that they are immune to error through misidentification. In this section, we briefly introduce this notion.

Wittgenstein (1958) distinguishes two uses of the first-person pronoun: a use as subject, exemplified in (22), and a use as object, exemplified in (23).

(22) I am in pain
(23) I have a bump on my forehead

The use of I as subject is immune to error through misidentification. Consider the statement Janet is in pain. There are several grounds on which this statement may be challenged. Among other things, someone may erroneously produce the statement Janet is in pain because she has mistakenly identified Janet as the person that is in pain when, in fact, it is someone else who is in pain. Consider now (22). Here the speaker cannot be wrong that it is she herself who is in pain. As Wittgenstein (1958, 66-67) puts it, discussing the sentence I have toothache, “there is no question of recognizing a person when I say I have toothache. To ask ‘are you sure it is you who has pain?’ would be nonsensical”.

However, there are also uses of the first person that are not immune to error through misidentification. Wittgenstein refers to them as uses of I as object. In the case of (23), for example, one can conceive of situations in which she knew that someone had a bump in her head, and legitimately wonders whether the one having a bump in her head was she herself.

It is worth mentioning that the notion of immunity to error through misidentification has rarely played any significant role in semantics. In fact, it is generally maintained that this notion only reflects the mode of access to the self that is associated with a given
class of circumstances of evaluation, rather than some property of the truth conditions associated with natural language expressions. The cases of de se that are discussed in the following section challenge this perspective.

6. Implicit de se

We are now in the position to substantiate the claim that the gerundive complements of remember and imagine are immune to error through misidentification. Consider the following inference (from Higginbotham 2003):

(24) I remember someone saying that John should finish his thesis by July
(25) As I am now assured, it was I who said it
(26) I remember saying John should finish his thesis by July

The inference is not sound: (24) does not follow from (25) and (26). My having acquired knowledge that it was me who said John should finish his thesis by July does not warrant the conclusion that I remember saying so. On the other hand, notice that (25) and (26) support the conclusion in (27).

(27) I remember myself/my saying John should finish his thesis by July

The facts illustrated in (24) - (27) show that (26) is incompatible with a scenario where the bearer of the memory is in doubt about the identity of the person who he remembers said John should finish his thesis by July. In fact, (28) seems contradictory, whereas (29) does not.

(28) Jim remembers saying that John should finish his thesis by July, although, now that he thinks about it, he is not sure it was he who he remembers saying it.
(29) Jim remembers himself saying that John should finish his thesis by July, although, now that he thinks about it, he is not sure it was he who he remembers saying it.

The intuitive reason for the contrast in (28) and (29) is that, when Jim remembers saying something to John he does not remember the proposition that he did so; rather, he remembers the subjective experience of saying something to John. It is in this sense that the control complement of remember is immune to error through misidentification relatively to the subject of remember.

The same conclusion holds for imagine. (30) sounds contradictory, whereas (31) does not. Also in this case, (30) reports Mary’s subjective experience of flying through space, which, as such, is immune to the possibility of misidentifying its experiencer, whereas (31) is also compatible with Mary’s imagining an individual, whom she identifies with herself, flying through space.
(30) Mary imagined flying through space, although she is not sure it was she who she imagined flying
(31) Mary imagined herself flying through space, although she is not sure it was she who she imagined flying

Following Recanati (2007) we will refer to the type of de se that is immune to error through misidentification on part of the bearer of the attitude as ‘implicit de se’. These examples also show that immunity to error through misidentification is a semantic property that enters, modulo embedding under some verbs of propositional attitudes, the compositional machinery that derives the truth conditions of complex expressions.

7. Acquaintance relations and implicit de se

How can we express the truth conditional import of immunity to error through misidentification? We argue here that the cases of implicit de se introduced above cannot be accounted for by relying on centered possible worlds and acquaintance relations.

According to Shoemaker (1968, 1994), immunity to error through misidentification emerges from the fact that the access to the self is immediate and non-observational, in the sense that it does not involve a perceptual or reflective act of consciousness:

«In introspective self-knowledge there is no room for an identification of oneself, and no need for information on which to base such an identification [...]. There are indeed cases of genuine perceptual knowledge in which awareness of oneself provides identification information, as when noting the features of the man I see in the mirror or on the television monitor tells me that he is myself. But there is no such role for awareness of oneself as an object to play in explaining my introspective knowledge that I am hungry, angry, or alarmed. This comes out in the fact that there is no possibility here of a misidentification; if I have my usual access to my hunger, there is no room for the thought “Someone is hungry all right, but is it me?”» (Shoemaker 1994, 257-258)

Shoemaker’s observation captures the intuitive link between being the immediate, introspective experiencer of a phenomenal experience and the impossibility to fail in identifying oneself as such. His observation also entails that the relation between the bearer of an attitude that is implicitly de se and the object of the attitude cannot be an acquaintance relation between two individuals, which is typically based on perception, memory, or communication. As such, acquaintance cannot be immune to misidentification. After all, one’s perception of oneself may be misleading, one’s memory of oneself imperfect, and what one has heard about oneself mistaken.

In what follows, we will propose that the implicit de se should not be explained in terms of a relation between individuals, but in terms
of relation between an individual and a novel type of object, a *phenomenal experience*.

8. Phenomenal consciousness

As Nagel (1974) famously put it, ‘there is something it is like to be a conscious organism’. That ‘something’ is our conscious expericer. It is inherently private and subjective to the point that, according to an important stream in Philosophy, it cannot be reduced to the physicalist vocabulary of science: ‘why is it that when our cognitive system engages in visual or auditory information processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C?’ (Chalmers 2010, 5).

The philosopher David Chalmers has proposed in a number of recent contributions that ‘a theory of consciousness requires the addition of something fundamental to our ontology’ (Chalmers 2010, 17). More specifically, Chalmers proposes that in order to establish a science of conscious phenomenal experience we must ‘take experience itself as a fundamental feature of the world’ (Chalmers 2010, 17). The approach to consciousness advocated by Chalmers qualifies as ‘dualist’, in that it assumes conscious experience as belonging to an independent ontological domain. It is proposed as a necessary step towards a science of consciousness, through the formulation of specific *bridging principles* to explain how experiences arise from physical processes.

It goes without saying that there is a rich philosophical debate surrounding these issues. However, our primary interest here is not to discuss whether dualism offers a correct description of the world (this, in fact, would be the task of the metaphysician). Our goal is rather to discuss the explanatory potential of dualism as a natural language metaphysics.

9. A dualist natural language metaphysics

Our hypothesis is that the world structure that is presupposed by natural language users reflects the dualist ontology proposed by Chalmers for the study of Conscious Phenomenal Experience. That is, we propose that the model for the evaluation of natural language expressions includes, on top of individuals, possible worlds, times, events, and any other class that is needed to provide a satisfactory account of speakers’ knowledge of meaning, *phenomenal experiences*.

Phenomenal experiences distinguish themselves from other objects for at least the two following properties. Firstly, experiences can be accessed by individuals only through *introspection*; the introspective access that an individual has to an experience is unique and private.
Let us refer to this property as the ‘Axiom of Privacy’ in the metaphysics of experiences. Secondly, the experiencer of an experience is automatically granted infallible knowledge of being the experiencer of that experience. This is so because, as discussed by Shoemaker, there is no process of identification of the experiencer with an object of the ‘external’ world by means of some sort of acquaintance relation or property ascription. Let us refer to this property as the ‘Axiom of Infallibility’ in the metaphysics of experiences.

Once a natural language metaphysics of this sort is assumed, the semantics of immunity to error through misidentification can be readily expressed in terms of reference to phenomenal experiences. Consider again sentence (13), repeated below as (32) for the reader’s convenience:

(32) Churchill remembers giving the speech

The sentence is a case of implicit de se, in the sense defined above. In a nutshell, our proposal is that remember denotes a relation between an individual, Churchill, and an experience. Accordingly, (14) is true iff, in the model of evaluation of the sentence, the set of experiences that correspond to Churchill’s memory includes the experience of giving the speech. Two observations follow immediately. Firstly, if it is the case that the only speech given in the relevant occasion was that of Churchill, only Churchill remembers the experience of giving it. This is because, given the Axiom of Privacy, in a model where Churchill is the individual having the experiential memory of giving the speech, no other individual in the model is having that same experience. Secondly, Churchill cannot be wrong in that it was him who gave the speech. Given the axiom of infallibility, Churchill, as the experiencer of the experience, is automatically granted infallible knowledge of being that experiencer.

10. Private language, other minds and Japanese evidentials

In the Philosophical Investigations (1953), Wittgenstein presents his well-known argument against ‘private language’. The fundamental purpose of the argument is to show that, under the assumption that meaning is reference, a language that refers to private objects, such as experiences and sensations, is logically impossible. Let us consider the argument by means of an analogy, by Wittgenstein himself: the ‘Beetle in a Box’. Let us imagine a community of individuals, each endowed with a box containing a beetle. Each individual can look at her own beetle, but only at her own beetle. That is, nobody can look at somebody else’s beetle. In such a context, people’s knowledge of what a beetle is depends solely on what each single individual sees in her
own box. Now, if the language of this community had a word *beetle* to refer to the type of object that is instantiated in each box, what would its meaning be? This is impossible to tell, submits Wittgenstein: It could be that each individual has a different thing in the box; it could be that the thing in the box is subject to constant change; it could even be that there is nothing in the box. We will never really know, because all we can know, as members of this (somewhat eccentric) community, is what is inside our own individual box. As it is already obvious to the reader, Wittgenstein’s analogy is with private sensations and experiences. To see an example, we cannot identify the meaning of *pain* with the private sensation of pain, because that would get us in the same troubles as with the beetle in the previous analogy. The conclusion drawn by Wittgenstein is that private language, a language whose meaning is defined in terms of reference to objects that are private, is logically impossible as it would be impossible to identify a consistent correlation between linguistic expressions and their referent.

The argument against private language threatens the hypothesis, defended in this contribution, that there are natural language expressions that refer to phenomenal experiences. However, its impact on the current proposal should be qualified in at least two respects. First of all, the argument assumes that the consistency of the rules of a language, crucially including the associations between expressions and their meanings, has to be, at least in principle, subject to public checking. This assumption has been vigorously criticized in several occasions (see, among others, Blackburn 1974-75 and Strawson 1954). In particular, it has been pointed out that the assumption corresponds to a form of verificationism that should not belong to a theory of linguistic meaning. Secondly, it is to be seen to which extent phenomenal experience is private. In fact, whereas it is true that direct access to phenomenal experience is available only to the experiencer itself, it is also true that individuals often engage in the practice of inferring other people’s mental states by relying on mechanisms of analogical induction and emphatic identification, a skill often referred to in cognitive psychology as Theory of Mind (see Apperly 2010 for an overview).

The latter observation correlates in an interesting way with a linguistic phenomenon observed in Japanese. Whereas the Japanese counterpart of the English sentence *I am in pain* is perfectly grammatical, the Japanese counterparts of *you are in pain* and *she is in pain* are judged as ungrammatical (Kuroda 1973). The descriptive generalization is that experiencer subjects can only occur in the first person in Japanese declarative sentences (Tenny 2006). Second and third person experiencer subjects can occur only when accompanied
by an evidentiality marker expressing the fact that the statement is made on the basis of 'indirect' evidence. That is, whereas a Japanese speaker may say of herself that she is in pain, she cannot say the same thing of somebody else. What she can say, however, is something that is roughly paraphrased as 'the speaker has indirect evidence that that individual is in pain'. These observations are of extreme interest in the light of the present proposal as they strongly suggest that natural languages may not only refer to experiences as a distinct class of objects but also overtly reflect actual properties of experiences, such as the fact that an individual can access another individuals phenomenal experience only indirectly.

Bibliographie
Delfitto D. and G. Fiorin. (in progress). De se readings as a window on the nature of Control. Université de Verona & Université de Utrecht.


