

Emergence of meaning, signals and the concept of consciousness

O surgimento de significado, sinais
e o conceito de consciência

Esteban Céspedes¹

ABSTRACT

Following an account of signaling games, one can show how meaning emerges and is preserved on the basis of the interactions between individuals and their environments. It is here argued that, as all concepts, a concept of consciousness is formed from a set of signaling games and is assigned a sense, from which its extensional reference can be postulated. It will be helpful to understand the contrast between what we may call a representationalist account of consciousness and an enactivist account. As argued, a consciousness state can be assumed and fixed by intensional reference. Thus, although the notion of consciousness may be explanatorily excluded, in principle, from a neurobiological language, it remains relevant in a semantic way. This is a consequence of what we may call the semantic gap between the mental and the physical.

Keywords: concept, sense, reference, enaction, mind, neurophenomenology.

RESUMO

De acordo com uma certa explicação dos jogos de sinalização, pode-se mostrar como o significado emerge e é preservado com base nas interações entre os indivíduos e seus ambientes. É aqui argumentado que, como todos os conceitos, um conceito de consciência é formado a partir de um conjunto de jogos de sinalização e é a ele atribuído um sentido, a partir do qual sua referência extensional pode ser postulada. Será útil entender o contraste entre o que podemos chamar de uma descrição representacionista da consciência e uma explicação enativista. Como argumentado, um estado de consciência pode ser assumido e fixado por referência intensional. Assim, embora a noção de consciência possa ser explicitamente excluída, em princípio, de uma linguagem neurobiológica, ela permanece relevante de maneira semântica. Isso é uma consequência do que podemos chamar de lacuna semântica entre o mental e o físico.

Palavras-chave: conceito, sentido, referência, enação, mente, neurofenomenologia.

¹ Valparaíso Complex Systems
Institute. Artillería 470, 2360448,
Valparaíso, Chile.
E-mail: estebancespedes@aol.com

Introduction

The enactive view of cognition, also called enactivism, is an account that should be understood as an alternative to representationalism (cf. Varela *et al.*, 1991; Fodor, 1998). According to representationalism, knowledge in general consists in a relation between a cognitive system and a previously given world. Cognitive systems receive impressions from the world and generate representations, which are manipulated in order to formulate symbolic expressions, other representations and mental maps of different sorts. The main types of expressions are sentences, which express propositions; names, which refer to objects; and predicates, which correspond to properties.

We may characterize a representationalist notion of the meaning of sentences as follows.

- *Representationalist meaning of a sentence.* The meaning of a sentence is a corresponding fact or set of facts and a sentence *s* is true, if it describes appropriately the corresponding facts, that is, if the predicates involved in *s* correspond in the right way to properties of the objects to which the names involved in *s* refer².

The notions of predicate meaning and of name meaning are implied in this characterization. We may understand the meaning of a name in terms of reference in the following way.

- *Representationalist meaning of a name.* The meaning of a name is the object to which the name refers.

And the notion of predicate meaning can be considered as follows.

- *Representationalist meaning of a predicate.* The meaning of a predicate *P* is a property of the object represented by the name of which *P* is predicated.

As an alternative to a representationalist account of meaning, the enactive view presents meaning as a characteristic that is produced by cognitive systems and their interactions with the environment. The world is not simply *represented* in cognitive processes, but also *enacted*, that is, brought forth, and cognitive processes do not consist merely in symbolic manipulation according to fixed rules, but arise thanks to the fact that mind, body and environment are in constant interaction. Thus, for instance, the process of a person understanding an expression does not simply involve a relation between a cognitive system and a previously given world, but is the result of the mentioned dynamic interaction and structural coupling.

According to Varela *et al.* (1991), mind and world arise simultaneously on the basis of enaction. How should we understand this arising? Which precise notion of meaning can be sustained on this basis? Which notion of world should we adopt? Could such a notion be compatible with a traditional notion of meaning? I will try to tackle these questions in the following sections.

As it will be shown, in order to understand what representation is, we need to characterize the notion of meaning, which is not only constituted on the basis of expressions and objects, but also on the basis of the field of phenomena, that is, of consciousness. Phenomenal experiences are at the roots of meaning formation. This is the main reason, I shall argue, why it seems hard to conceptualize and explain them. It will be proposed that although descriptions about conscious states can be explained in more fundamental terms, the intentional content of those states cannot.

The paper's structure is the following. In the next section, I will introduce a general notion of emergence and the signaling games account. According to this account, meaning may arise as an emergent property from a system of signals. The third and fourth sections explore how signaling games may be applied to neurophenomenology to provide an account on the notion of consciousness. Then I will propose a definition of the notion of concept, based on the distinction between sense and reference, as well as on the signaling games account. This should provide a framework that clarifies in which sense consciousness is reducible and in which sense it is not, as well as how different concepts of consciousness may be integrated according to an enactivist and neurophenomenological approach.

Emergence and signaling games

In order to grasp the enactivist idea that mind and world arise together, the notion of an emergent property may be of great help. According to a general characterization, emergent properties are properties that arise after the interactions between the basic elements of a given system reach a determined degree of complexity (cf. El-Hani and Pereira, 2000; Butterfield, 2011). In this sense, novelty is one of the main aspects of emergent properties. The emergent properties of a system are properties that the system did not have before they arose.

Robustness is also an important aspect of many emergent properties: They have a certain structure that persists during a considerable period of time. Another relevant and recurrently discussed aspect of emergent properties is their irreducibility. In general, a description of an emergent property cannot be reduced to any set of descriptions of the basic

² Regarding truth, we should also consider cases in which two sentences with different meanings are made true by the same fact. Imagine the fact that it is raining. Both sentences, "The sky is cloudy" and "The air humidity is high", are made true by that same fact. So, a sentence can also be true if it is implied by a sentence that describes some fact appropriately. However, according to a representationalist notion, the meaning of "The air humidity is high" is not the fact that it is raining, but the fact that the air humidity is high.

constituents of the system in which it arises. The observation of the instantiation of an emergent property can be called *emergent phenomenon* and the process in which such a property arises, *emergence*.

I will argue in what follows that the meaning of a given expression arises as an emergent property of a system of interactions. Which are the constituent elements in a system from which meaning emerges? Which are the relevant interactions involved? The notion of a signal may shed some light on these issues.

Signaling games serve as models on the basis of which different types of linguistic and behavioral phenomena might be explained (cf. Shannon, 1948; Lewis, 1969; Skyrms, 2010). A signal can be simply understood as a process that is able to influence other processes causally. Light beams, waving hands and sounds are good examples thereof. Sentences can also be considered as signals. Predicates and names can be considered as parts of sentences, but also, when uttered in isolation, as one-word-sentences. Thus, for instance, the expression “spider” might mean “There is a spider!” and “cold” might mean “The water is too cold.” However, as it should become clear in a moment, signals do not carry meaning intrinsically. Before explaining this, let us focus on the basic features of a signaling game.

A basic signaling game consists of two players interacting with each other: the sender and the receiver. After observing a state, the sender emits a signal to the receiver, who does not have any previous information about the state. After getting the signal, the receiver decides to perform an act, which may affect both. If the act performed by the receiver is coherent regarding the state observed by the sender, both will get a positive pay-off. If it is not, the pay-off will be zero. In this sense, we might say that sender and receiver have a common interest in communicating successfully.

Now, at which point is a signal taken as a piece of meaningful information? In order to answer this, one should consider the development of a signaling game. Note that, after the sender and the receiver get their pay-offs, they may be confronted with a new situation of the same kind, that is, the sender may observe the same type of state and send the same signal type to the receiver, after which the receiver could perform the same kind of act that she performed before. Of course, it may also be that, in a new situation, the sender observes the same type of state, but decides to send another kind of signal. After that, the receiver may also try something new. If both get a positive pay-off in such a combination, they might decide to maintain that strategy. If with a given combination their pay-offs are zero, they might choose to change their strategies, that is, the sender may decide to send a different signal or the receiver may decide to perform a distinct act. The main point is that learning is guaranteed.

In a signaling game sufficiently developed, the meaning of a signal emerges after the number of interactions between the sender and the receiver in which that signal is involved increases considerably. At some point, because of several reinforcements, the sender stays with the strategy of sending a particular

type of signal after observing a given state type and the receiver stays with the strategy of performing a particular act type after receiving that particular type of signal. We may describe this by considering how, from the perspective of the receiver, signals may move the probabilities assigned to the occurrence of a state. Thus, at the beginning of a signaling game, the probability assigned by the receiver to a state of a two-member partition, after receiving a certain signal, should be around .5. After a while and thanks to several reinforcements, the receiver might assign a higher probability to the occurrence of that state type, given the same signal type. The informational quantity, I , carried by a signal s about a state F can be characterized in terms of probabilities as follows (cf. Skyrms, 2010, p. 36):

$$I_F(s) = \log_2 [p(F|s) / p(F)]$$

Applying the logarithm to the definition permits to characterize informational bits. Thus, when the probability of the occurrence of a state F , given that the signal s is received, is not different from the prior probability of the occurrence of F , s is considered as carrying no information regarding that state. Now, suppose that a receiver is only focused on two possible states and regards them equiprobable before receiving the signal s . If s moves the probability of one of those states to 1 or to 0, it contains one bit of information.

Signals also carry information with regard to the observed acts performed as reactions to signals. The quantity of information about an act that may be carried by a signal can be characterized as follows, considering an act type A (cf. Skyrms, 2010, p. 39):

$$I_A(s) = \log_2 [p(A|s) / p(A)]$$

As already mentioned, signals do not carry information intrinsically. It may be that at the beginning of a signaling game, a fully new signal type s contains no information at all. If, during some interval of the game's development, the information about a state F carried by s has increased up to one bit or more and been maintained so, we may say that a meaning of s has emerged. (When this occurs, the sense and the reference of that signal can also be determined, as we will see later.) For the involved receiver, the signal s may mean at that point that F is a fact. Note that this notion of meaning is compatible with enactivism. The meaning of s is brought forth by the interactions between the receiver, the sender and the environment. The signal s is not simply the representation, for a receiver, of the state F . It is a feature of a system in which the receiver is embedded as a constituent and fundamental part.

As an example, consider vervet monkeys. They have developed different alarms for different kinds of predators (Seyfarth *et al.*, 1980; Skyrms, 2010). For instance, when a leopard is near, some individuals start barking, thereby alerting their partners, who generally react appropriately running up a tree. When an eagle is approaching, the call is different. Some individuals start coughing, to which their partners react by hiding in the

bushes. It is more than plausible to think that such a communication system emerged from several iterations of inter-generational signaling games. We could say that the expression of coughing means that an eagle is near. Such a meaning is a feature that emerged from the interactions between many individuals in many situations involving eagles and other flying predators.

The signaling games account of information is similar in some way to teleological theories of meaning, also called teleosemantic theories. Fred Dretske's indicator semantics is one example. According to Dretske (1981), a system represents some property just in case it has the function of indicating that property, that is, of providing information about it. Another well-known example of a teleosemantic theory is Ruth Millikan's (1984) account based on the notions of consumer and benefit. According to this account, the content of a representation depends on how consumers of that representation are benefited and able to co-adapt with its producers. David Papineau (1984) developed a teleosemantic account that explains representations in terms of desires. Desires, together with beliefs or representations, can produce actions. Considering this, the content of a representation r depends on how a desire d is satisfied by an action based on r and d . Note that the notions of indication, benefit and desire express intention. Skyrms (2010, p. 43) argues that no sort of intentionality is needed to determine informational content. However, appropriate models involving the notion of intentionality can be added to the theory of signaling games without losing coherence.

Although the notion of intentionality may not be necessary to account for informational content, it may be necessary to account for what concepts, meanings and representations are. In the following sections, I will try to show briefly how this could be done, having in mind the main aspects of the enactivist perspective and focusing on a particular kind of intentionality: consciousness. As mentioned in the previous section, phenomenal experience grounds concept formation. And this kind of experience is essentially intentional. If we want to claim that an agent possesses a given concept, we have to assume that it is able to have some kind of phenomenal experience that is, under some respect, similar to our phenomenal experiences regarding the same set of entities. The present proposal is partly about taking this into account. Not doing this may lead to confusions not only in trying to conceptualize meaning but also consciousness. Now, before focusing on the intentionality of consciousness, I will briefly introduce in the following section the neurophenomenological account, which implies, as argued later, different meanings of "consciousness." This grounds the idea that the term "consciousness" may correspond to different things, depending on different sorts of intentionality.

Neurophenomenology, signals and meaning

The issue of explaining how the world of physical processes is accompanied by consciousness is well known as the

hard problem of consciousness (Chalmers, 1995). According to the research program of neurophenomenology, which was proposed by Francisco Varela (1996) and is in some sense an extension of enactivism, not only neurobiological descriptions, but also descriptions concerning the lived and embodied experience should be taken into account if one wants to deliver a satisfactory answer to that problem. The hard problem of consciousness should be tackled by an account in which the notions of subjective experience and meaning play central roles.

As David Chalmers (1997) claims, neurophenomenology seems to be on the right track concerning the hard problem of consciousness. Now, it may be important to consider that Varela diminished the idea that the notions of signal and information could contribute considerably to the debate. However, I think that these notions can in fact contribute, if the notion of meaning, as defined above, is applied carefully. One may think that the notion of a signal, understood as a non-symmetrical process by which one event influences another, is not compatible with the notion of co-dependence that is fundamental for the enactivist perspective. As a response, we should note that causal influence based on signals may be compatible with co-dependence, if both kinds of influence occur at different levels. For instance, species and the features of the environment associated with nutrition can be considered as co-determined at the evolutionary level, but processes by which a single organism gets nutrients from the environment can be (and may be better) described in terms of unidirectional signals. Note that the basic notions of a theory of signals are not enough to tackle the problem of consciousness in a way that follows the main aims of neurophenomenology faithfully. One could perfectly defend a representationalist account of cognition in terms of signals, assuming some sort of isomorphism between the environment and the mechanisms involved in the emission and interpretation of signals. Nonetheless, characterizing meaning in terms of signals may help us to account for consciousness as an embedded feature of a signaling system as well.

There are two main kinds of explanations in which signaling games might be applied relevantly within an account of consciousness. One kind of explanation seeks to describe how certain brain patterns emerge from the interactions in neural networks and are correlated with conscious mental processes. We may call any explanation of this sort a *neural correlate explanation of consciousness* (cf. Koch, 2004). The main problem with this sort of explanation is the fact that to provide descriptions about correlations between neural activity and reported (or assumed) conscious experiences does not seem to be enough to explain how consciousness emerges together with neurobiological processes (cf. Chalmers, 1995; Noë and Thompson, 2004). This is known as the explanatory gap between mental and physical processes (Levine, 1983). Related to this weakness, a second sort of explanation, which is also grounded on the ideas of signaling games, is crucial. According to it, one may provide descriptions involving the interactions between an organism and its environment in order

to explain how the concept of its conscious mental processes is shaped during signaling games. We may call any explanation of this sort *enactive explanation of consciousness*. This kind of explanation may involve factors at different levels, such as evolutionary, ecological, physiological and neural factors.

Usually, the explanatory gap between neurobiological descriptions and descriptions about conscious states appears in contexts in which neural correlate explanations of consciousness are intended. In these contexts, it seems very difficult to explain how some physical processes are accompanied by conscious processes. Considered as mere correlates, conscious states and neurophysiological processes are not linked in a way according to which we may satisfactorily explain the former on the grounds of the latter. In principle, there is no apparent incoherence in conceiving physical replicas of conscious human persons that do not have conscious experiences (cf. Stout, 1931; Block, 1980; Chalmers, 1996). Anyhow, the fact is that we usually assume that certain kinds of behavior are accompanied by conscious states. Enactive explanations may contribute to understand this.

There is a worry about grounding an account of meaning on the notions of signals and information. As a formal notion and as defined above, it seems that information is neither related to causation nor to a concept of truth. How can meaning, in a representationalist sense, emerge from causal interactions? There are mainly two notions of causation playing a role here. On the one hand, physical causation is important, of course (Dowe, 2000). We assume that senders, receivers and observed states are connected by physical processes. On the other hand, causation as difference-making is also crucial (cf. Ney, 2009). The information associated with a signal expresses how it makes a difference in an agent's representations. In this sense, representations depend causally on the occurrence of expressions and on the context of their occurrence. Additionally, we have to assume that signals depend causally on certain states of the world. Of course, here the notion of world can be interpreted following enactivism. When the information of a certain state type contained in a signal is high and stable, the description of such a state may be taken as a representation and the state may be taken as a fact. In these cases, representations can correspond to facts in a representationalist sense. Then, we may say that a signal (or a set of signals) is a true proposition if it describes correctly the fact with which it is associated, i.e. if it provides the highest informational value with regard to that fact. Clearly, the concept of information is not enough to understand this. A broader framework is needed. I will try to explore in the last section the general aspects on which such a framework could be based. Before that, let us turn again to the notion of consciousness. I will argue that, to be in a state of consciousness and to attribute consciousness to other agents, one has

to possess some notion of the state of which one or the other agent is conscious. Thus, to be aware of one's own state of consciousness, as well as to attribute consciousness, one needs to understand the concept of consciousness and how it is used. The following section is concerned with the main features of such a concept.

Consciousness and concepts

Consider the following case. Suppose, for instance, that John sees a black stain on the wall. Can he be mistaken about that? For sure, on the basis of a posterior perspective, it is not required that John is right about what he thinks he sees now. It may be that, after reporting that he sees a black stain on the wall, he realizes that it was actually a fly on the wall what appeared to be just a black stain. In that case, we might say that he was wrong, on the basis of his new experience. However, we cannot claim that, from the perspective of his experience, he was wrong when he thought he was seeing a black stain on the wall and was aware of that. The fact that it seemed to John that he saw a black stain on the wall must have been correct in some sense (if that was actually his experience). A state of consciousness or awareness is a state of that sort and descriptions thereof cannot be corrected on the basis of new experience.

The description of a state of awareness is neither a description that one can correct on the basis of external experience. We might say, for instance, on the basis of a very specific description of John's recent neurobiological states, that John is wrong when he says that he is seeing a black stain on the wall. We may have a great theory about how visual space is generated and describe some particular visual space based on hue, brightness and saturation. Then we may claim that such a space is actually what John is perceiving. Nevertheless, if he reports that it seems to him that he is seeing a black stain on the wall, it must be that way in some relevant sense. A state of awareness is a state of that sort and not a state postulated from some external perspective.

The characterization just made refers to what we may call *consciousness as experience* or as what it is like to have some particular experience. This should be distinguished from what we can call *conscious experience*³. Consciousness, in this latter sense, is a kind of mental state about another mental state (cf. Rosenthal, 1986; Carruthers, 2000). If John is aware that he is seeing a black stain on the wall, he is in a mental state of awareness *about* his seeing. Also, we might say that, at least for some species, awareness is nothing but self-awareness or, perhaps more cautiously, that, in some cases, there is no state of awareness without self-awareness (Damasio, 2000; Gallagher, 2000). John is aware about *his own* state of seeing. Related

³ This distinction has the same general basis as Ned Block's (1995) famous distinction between *phenomenal consciousness* and *access consciousness*. However, the characterizations of the concepts involved in my distinction might be conceived as crucially different from his characterizations.

to this, we may further claim that although the so-called first-person data is privileged data, it is not data of a special epistemic kind. Thus, the distinction between first-person and third-person data might be misleading.

To be in a mental state of self-awareness or of conscious experience, one has to possess a concept of the mental state of which one is conscious⁴. So, structurally, the awareness of perceptual experience is very similar to the use of an expression learned in a signaling game. For example, it does not make sense to say that a vervet monkey is trying to communicate to his partners that a big bird is approaching without also assuming that he has some notion of a big bird (or, more generally, of some flying, dangerous thing). As well, we would not think that John was aware of seeing a black stain on the wall without assuming that he has some concepts of seeing, of the color black and of a stain. Considering this similarity, conscious experience—just as the capacity of grasping truths and meanings—cannot be fully understood without considering a subject's previous experiences within a set of signaling games. This is related to the already mentioned thesis that processes of concept formation based on signaling games involve phenomenal experiences as essential, constitutive parts.

We may now consider the following characterization of conscious experience. Let *A* and *M* be two mental states. If *A* is a state of awareness about state *M*, we shall call *M* the phenomenon or experience. A phenomenon is not something that we can describe in detail, although we may describe *A* on the basis of concepts that are about *M* (Figure 1).

While states of conscious experience can be described in neurobiological terms, one cannot describe in those terms the particular mental state on which a state of consciousness is (or seems to be) focused. It would be like trying to construct a neurobiological description of the black stain that John reports he is seeing. We cannot. We can only describe in neurobiological terms his state of awareness, but not what he feels he sees.

How should we understand these aspects of (the concepts of) consciousness? The account of signaling games, introduced in the second section, will be important to tackle this question. According to it, meanings are stable states within signaling games. Of course, we should not only consider the meaning of consciousness, but also the meanings of the mental states that we want to attribute. Another crucial point is, as I have tried to show, that phenomenal experience must be considered as a key aspect of any signaling game that is involved in meaning formation. This allows us to propose an enactivist notion of consciousness that can be assumed within the framework of neurophenomenology. As explained, neurophenomenology suggests that we should take into account

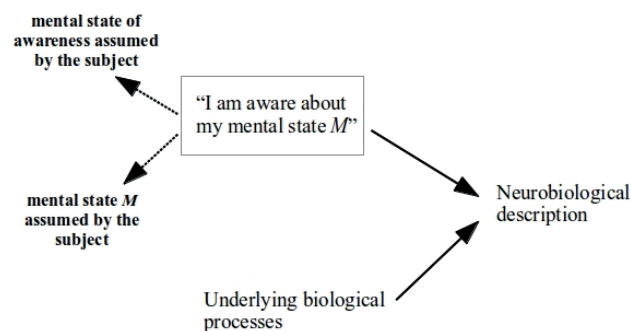


Figure 1. A subject's report about his mental state may involve both a focus on that state and on the state of self-awareness. Neurobiologists are able to compare his report with some biological basis, but they may not be able to describe the particular phenomena on which the subject is focused.

neurobiological as well as phenomenological descriptions if we want to arrive at an appropriate theory of consciousness. In order to develop such a theory, we have to consider the different aspects of concepts. I suggest in the following section that such aspects can be characterized on the basis of the classical distinction between sense and reference, together with distinctions between different notions of reference. Consequently, this should also help us understand different aspects of the concept of consciousness.

Dividing the world and putting worlds together

As argued, one can only be aware of one's own conscious state if one already possesses some notion of consciousness. I have focused on the distinction between states of consciousness that are describable in neurobiological terms and the ones that aren't. This distinction not only implies that there may be different notions of consciousness, but also emphasizes different aspects that constitute any of those particular notions, namely extensional and intensional aspects. On the basis of these distinctions, I will try to elaborate further on the thesis, mentioned in the previous section, that although descriptions about conscious experience are explainable in terms of neurobiological descriptions, there is a point of view according to which the content of conscious states cannot be thus explained.

Considering the signaling games account again, if we want to attribute consciousness states to an agent, we can neither focus just on the agent's behavior, including verbal reports, nor on its physiological states. We have to consider

⁴ One could consider the following objection: Some non-human animals seem to have conscious experiences, but they are not involved in the creation or transmission of meanings. My general claim is that any individual that can be, according to our language, in a mental state of conscious experience must also have a concept of the particular mental state of which it is aware. Now, with regard to the creation and transmission of meanings, the notions of concept and meaning considered in this work are broad enough to be applied to non-human animals.

the agent's concepts of the qualia that we want to attribute to it, our concepts of those qualia, as well as our concept of consciousness, of course⁵. Signaling games allow us to describe these concepts within a framework that is compatible with enactivism and neurophenomenology. But how can we say, for example, that John is aware that he is seeing a dot on the wall if we neither have direct access to his experience nor to the dot's material constitution itself?

An idea that may be very helpful, following Kant, is the distinction between two meanings of the notion of world. On the one hand, the world can be understood as a thing-in-itself, an entity that is completely independent from the knowing subject. It cannot be conceptualized in depth, and one of the few features that one may attribute to it is that it is the ground of our perceptions. On the other hand, the world is the set of all things as known empirically. On this basis, it can be described theoretically. Sets that we construct based on considerations of the world as a thing-in-itself can be defined by their intensions (i.e. the properties to which we can point directly), but not by their extensions. Sets constructed from a theoretical ground can be called ontologies and must, at least partly, be defined on the basis of their extensions. Both meanings of the idea of world also express two ways of considering a concept. So, in order to explain how we can think about the world in those two ways, we must explain what parts within the structure of a concept permit us to make such a distinction. On the basis of the notion of a signaling game (and taking some distance from a Kantian framework on this matter), we may characterize a concept as follows:

Concept. For a subject *S*, involved as a sender, receiver or observer of a signaling game, a signal type *c* is considered to be a concept if and only if the following conditions hold:

- (a) According to *S*, *c* has a high and stable informational content associated with either a state type or an act type.
- (b) *S* assigns *c* an intensional reference, i.e. a relation that links *c* to some relevant state or act type.
- (c) *S* assigns *c* a sense, i.e. a role that *c* plays in some relevant set of sentences (which may describe signaling games).
- (d) *S* assigns *c* an extensional reference, which is an entity of an ontology brought forth by *S*.

Following a traditional distinction proposed by Gottlob Frege (1892), the meaning of an expression can be determined, on the one part, by a sense, that is, by the role that the

expression plays within a set of representations, and, on the other, by a reference, that is, the relation that connects the expression with an assumed independent reality. Note how we can distinguish two sorts of reference: intensional reference and extensional reference (cf. Hesse, 1974). On the basis of intensional reference we can talk about the world-in-itself understood as the ground of our perceptions (without describing it), while on the basis of extensional reference we can put things in defined categories and describe them in detail⁶.

An ontology can be constructed based on the extensional references of different expressions involved in some relevant set. We may say that a world is brought forth (or enacted) when such an ontology is assumed either pragmatically or theoretically. Note that we can still account for the notion of representation in a traditional way: Expressions represent entities within an ontology. However, we do not have to understand sense and reference as representationalism does, neither must we assume that cognition is based on representation in a strong sense. According to the notion of a concept just characterized, entities of an ontology are brought forth in enactivist fashion, that is, they emerge from signaling games, from interaction networks involving bodies, minds and environments.

It is important to note that extensional reference cannot be fixed without fixing intensional reference. Semantically, the idea of other agent's phenomenal experiences plays a similar role as the Kantian notion of world-in-itself. We do not have access to those experiences, we cannot provide extensional characterizations of them, but we can point at them, think of them. After pointing at them, we can, on the basis of our ontology, define expressions referring to them. The same kind of intensional reference to phenomenal experience is needed to understand concept formation in general. The latter is a thesis that I have mentioned and defended throughout this work. We cannot attribute concept possession without attributing (by assumption, at least) subjective experience and we cannot understand subjective experience without understanding this kind of reference (Figure 2).

We may also account for the enactivist idea that world and mind emerge together. Roughly, a mind can be individuated on the basis of its conceptual abilities. Such abilities to form and use concepts, as well as the entities to which those concepts refer, emerge together within signaling games, i.e. systems of interacting individuals and environments.

Consider this further simple example: Suppose that Julia feels pain after biting her tongue. Her concept of pain is constructed on the basis of types of perceptual experience, particular experiences and information acquired by other sources⁷.

⁵ A huge effort is being made to improve introspection methods and to compare them with neurophysiological data (cf. Froese et al., 2011). Comparisons of that kind will not clarify the most crucial questions about consciousness if they are not included in a research agenda that takes semantic and ontological issues seriously.

⁶ Although experiences are key elements of concept formation, the notion of phenomenal experience does not have to be included in the analytic definition of the notion of concept given here. It should, however, be included in a generative description of concept formation.

⁷ Consider the following possible objection: The newborn can feel pain without having a concept of pain. I agree. But this should not be a problem, accepting the simple assumption that we are the ones that need a concept of pain in order to express that babies can feel pain. Furthermore, I assume that Julia's brain is developed enough to produce states of consciousness, as they were characterized.

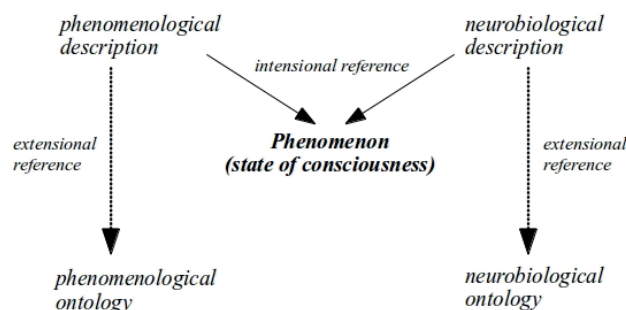


Figure 2. Even if phenomenological and neurobiological descriptions can refer to the same phenomenon, the ontologies derived by them may be fundamentally distinct.

She may refer to particular experiences of pain in a direct way, without describing them. Also, what she has learned allows her to associate different types of perceptual experience, such as getting sunburned or cutting one's finger with a knife. Qualitatively, those perceptual experiences are similar in a relevant way. Now, their ground is, Julia may assume, some distinct set of states. For instance, she may assume particular neurobiological mechanisms that produce her pain or that her experience of feeling pain is based on some more basic state, the experience of pain itself.

The sense of the concept of pain is, for Julia, determined by the role that it plays regarding her beliefs and actions. Thus, in contexts that include her experiences of pain, the sense of the concept of pain must not be associated with the sense that it may have on the basis of a set of signaling games constituted by, say, conversations with neuroscientists. And the reference of the concept of pain may be, for her, a property postulated within an ontology that satisfies the conditions related to the concept's sense. So, if she is focused on some sort of introspective set of descriptions, the reference may be a phenomenon type or, in some cases, a particular kind of mental state. Now, if Julia was focused on a scientific context, the sense of the notion of pain would be different accordingly. In such a case, she may postulate that the reference of the term "pain" is some neurobiological mechanism or some complex set of neural processes.

Naturally, concepts can also be grasped by individuals that observe a signaling game without participating in it. Suppose that John has been observing Julia, focused on what might be her concept of pain. From John's point of view, Julia's concept of pain must be constructed on the basis of Julia's behavior, as well as on John's notion of pain. When Julia reports to John that she feels pain, John may think of Julia's past reported experiences, as well as of his theoretical knowledge regarding pain. He can associate her report to any stipulated, non-behavioral state. Whatever that state may be, John cannot grasp it, but just assume it.

On the basis of John's observations, the sense of Julia's concept of pain may be determined by the role her behavior plays in certain situations that he associates with pain.

He may also study how her reports are correlated with her neural states. In that case, the sense of the term "pain" could be based on the role that it plays within neurobiological theories of pain sensation. The term's reference, i.e. the set of entities of a postulated ontology to which it corresponds, would vary depending on its sense.

Based on an ontology assumed from John's perspective, Julia's phenomenal field can also involve entities such as inner pictures and mental images, which cannot serve to specify a pure phenomenon extensionally. And on the basis of Julia's assumptions, those kinds of entities may also have relevance. A subject's own internal world is just as inaccessible as for an external observer. In cases of introspection or conscious experience, a subject is focused on her own behavior. Statements like "I saw a black stain on the wall" or "I feel pain" can be understood as behavioral reports. A subject may also think about the ground of her phenomenal experience, her internal world, but could not describe it as it is merely by introspection. Introspection could only help her to learn about how she represents her experiences. Experiences themselves, so to say, cannot be grasped at all.

A crucial question related to the hard problem of consciousness is the following: Are all descriptions based on phenomenal ontologies (e.g., ontologies of folk psychology or ontologies related to phenomenological reports) explainable in terms of physical or neurobiological descriptions? For instance, are all descriptions about pain explainable by descriptions about neural states? On the basis of the notion of meaning considered here, the answer is yes. In principle, we cannot exclude the possibility that a general theory about pain may be able to put our phenomenon-based and our neurobiology-based ontologies together. However, the hard problem may not be solved that way.

There is a gap regarding meaning that seems to be at the basis of the explanatory gap between consciousness and neuroscience, a *semantic gap* between the concepts of consciousness and of physical processes. It can be considered in a general way as follows. Let n be some neurobiological description and q be some phenomenological report or description. Both are about a particular, assumed phenomenon Φ . The relation that links n and q to Φ is intensional reference and is what constitutes the semantic gap. The thought of n and q that grounds such a relation can be considered to be a *de re* thought, i.e., a direct and purely referential thought (cf. Jeshion, 2010). The descriptions n and q , as well as the concepts involved in them, also refer to particular entities and their properties, which are part of the ontologies associated with those descriptions. The relation that links n and q to those entities is extensional reference.

Consider the sentence "Julia feels pain" and the concept of feeling pain that occurs in it. The intensional reference of that concept is Julia's supposed experience or, in other words, whatever her feeling is about. Note that neither the reference of the proposition expressed by that sentence nor the reference of the term "pain" that appears in it must be fully

determined extensionally. With regard to intensional reference, we may say nothing specific about what it is or how it is categorized; we just point at it. At the same time, we postulate it as something real, as a thing-in-itself. By contrast, the sense of the sentence “Julia feels pain” would be the role that it plays within a set of sentences, a set of theories or within a conceptual network. Suppose that such a network involves some neurobiological model of pain. An ontology may be constructed on that basis, which permits to determine the extensional reference of the concept of pain.

Now, although the ontologies associated with the neurobiological description n and the phenomenal description q could be unified at a given point within theoretical development, consciousness, constructed on the basis of intensional reference, must be still assumed as a relevant ground and understood as separated from those ontologies⁸. This shows how, from a semantic point of view, the neurobiological world has to be accompanied by consciousness, even if the term “consciousness” could be described by—or even end up being completely excluded from—our best neurobiological ontology⁹.

Conclusions

In this article, an enactivist notion of meaning based on the account of signaling games has been proposed. As is well known, the hard problem of consciousness is, roughly, the question of why and how conscious states arise from and are accompanied by physical processes. As a way of tackling this problem, neurophenomenology proposes that phenomenological and neurobiological accounts should be articulated and not be considered as independent disciplines. Since neurophenomenology must be understood as an extension of the enactive account of cognition, one may try to provide a notion of meaning developed on the basis of enactivism that could also be applied to the questions that neurophenomenology seeks to answer. This was a main aim of this work.

Concepts, as proposed here, are types of signals that have reached a high and stable informational content because of the numerous interactions involved between the members of the system in which they have been instantiated. When the informational content of a signal is high and stable enough, individuals may assign to it a sense. We do not need any fundamental notion of intentionality to determine a signal's informational content. However, intentionality is needed to give an account of what a concept is, from the perspective of a subject. The sense of a concept is determined by the role it plays within a set of signaling games. On this basis, one may project

the concept's extensional reference into an ontology. Individuals also assign an intensional reference to concepts, which is a direct relation between the concept and some assumed set that does not have determined structure. Sometimes, such a set is called inner world, sometimes external world. Conscious states, one may postulate, are parts of inner worlds.

Considering this, the problem of whether consciousness could be explained in terms of neurobiological descriptions can be given two radically distinct interpretations. The first produces the following formulation: Are all descriptions about conscious states reducible, in principle, to some set of neurobiological descriptions? This is an empirical question and its answer is very likely to be yes. The second way of formulating the problem is this: Is there a possible set of neurobiological descriptions that could provide an explanation of any description about conscious states and, as well, be able to eliminate the concept of consciousness? On the semantic grounds explored in this work, the answer is no.

Acknowledgements

A previous version of this article was presented in Valparaíso at the seminar *Neurophenomenology: Building a Science of Experience*, during the summer of 2016. I thank the organizers and the participants. The paper started to take form thanks to the discussions we had then. I am also deeply indebted to Alan Baker, Peter Baumann, Richard Eldridge and Krista Thomason for their valuable comments, clarifications and criticisms on a more recent version of the paper during a very enriching discussion at Swarthmore College. Additionally, this work was financially supported by the Chilean Commission for Scientific and Technological Research (CONICYT/FONDECYT post-doctoral fellowship/project No: 3160180).

References

- BLOCK, N. 1980. Are Absent Qualia Impossible? *Philosophical Review*, **89**(5):257-274. <https://doi.org/10.2307/2184650>
- BLOCK, N. 1995. On a Confusion about a Function of Consciousness. *Brain and Behavioral Sciences*, **18**(2):227-247. <https://doi.org/10.1017/S0140525X00038188>
- BUTTERFIELD, J. 2011. Emergence, Reduction and Supervenience: A Varied Landscape. *Foundations of Physics*, **41**(6):920-959. <https://doi.org/10.1007/s10701-011-9549-0>
- CARRUTHERS, P. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge, Cambridge University Press, 368 p. <https://doi.org/10.1017/CBO9780511487491>

⁸ In this sense, Varela's program of neurophenomenology could contribute to the debate by providing better integrative explanations of reports about subjective experience and of neurobiological descriptions. In other words, our understanding of different enacted concepts of consciousness can be improved notably on its basis. However, neurophenomenology cannot contribute to definitively bridge the semantic gap between states of consciousness and neural processes, not only because of its methodology, but because of the very nature of the problem.

⁹ Interestingly, the considered semantic gap somehow implies the explanatory gap but, as well, a permanent ontological gap.

- CHALMERS, D. 1995. Facing up to the Problem of Consciousness. *Journal of Consciousness Studies*, **2**(3):200-219.
- CHALMERS, D. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford, Oxford University Press, 432 p.
- CHALMERS, D. 1997. Moving forward on the Problem of Consciousness. *Journal of Consciousness Studies*, **4**(1):3-46.
- DAMASIO, A. 2000. A Neurobiology for Consciousness. In: T. METZINGER (ed.), *Neural Correlates of Consciousness*. Cambridge, MIT Press, p. 111-120.
- DOWE, P. 2000. *Physical Causation*. Cambridge, Cambridge University Press, 236 p.
<https://doi.org/10.1017/CBO9780511570650>
- DRETSKE, F. 1981. *Knowledge and the Flow of Information*. Cambridge, MIT Press, 273 p.
- EL-HANI, C.; PEREIRA, A. 2000. Higher-level Descriptions: Why Should We Preserve Them. In: P.B. ANDERSEN et al. (eds.), *Downward Causation*. Aarhus, University of Aarhus Press, p. 118-142.
- FODOR, J. 1998. *Concepts: Where Cognitive Science Went wrong*. Oxford, Oxford University Press, 192 p.
<https://doi.org/10.1093/0198236360.001.0001>
- FREGE, G. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, **100**(1):25-50.
- FROESE, T.; GOULD, C.; BARRETT, A. 2011. Re-Viewing from Within: A Commentary on First- and Second-Person Methods in the Science of Consciousness. *Constructivist Foundations*, **6**(2):254-269.
- GALLAGHER, S. 2000. Philosophical Conceptions of the Self. *Trends in Cognitive Sciences*, **4**(1):14-21.
[https://doi.org/10.1016/S1364-6613\(99\)01417-5](https://doi.org/10.1016/S1364-6613(99)01417-5)
- HESSE, M. 1974. *The Structure of Scientific Inference*. New York, Macmillan, 309 p.
- JESHION, R. (ed.). 2010. *New Essays on Singular Thought*. Oxford, Oxford University Press, 256 p.
<https://doi.org/10.1093/acprof:oso/9780199567881.001.0001>
- KOCH, C. 2004. *The Quest for Consciousness*. Placerville, Roberts and Company, 429 p.
- LEVINE, J. 1983. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly*, October, p. 354-361.
- LEWIS, D. 1969. *Convention: A Philosophical Study*. Cambridge, Harvard University Press, 228 p.
- MILLIKAN, R. 1984. *Language, Thought and Other Biological Categories*. Cambridge, MIT Press, 367 p.
- NEY, A. 2009. Physical Causation and Difference-making. *British Journal for the Philosophy of Science*, **60**:737-764.
<https://doi.org/10.1093/bjps/axp037>
- NOË, A.; THOMPSON, E. 2004. Are there Neural Correlates of Consciousness? *Journal of Consciousness Studies*, **11**(1):3-28.
- PAPINEAU, D. 1984. Representation and Explanation. *Philosophy of Science*, **51**(4):550-572.
<https://doi.org/10.1086/289205>
- ROSENTHAL, D. 1986. Two Concepts of Consciousness. *Philosophical Studies*, **49**(3):329-359.
<https://doi.org/10.1007/BF00355521>
- SHANNON, C. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*, **27**(3):379-423; **27**(4):623-656.
- SEYFARTH, R.; CHENEY, D.; MARLER, P. 1980. Vervet Monkey Alarm Calls: Semantic Communication in a Free-ranging Primate. *Animal Behaviour*, **28**(4):1070-1094.
[https://doi.org/10.1016/S0003-3472\(80\)80097-2](https://doi.org/10.1016/S0003-3472(80)80097-2)
- SKYRMS, B. 2010. *Signals: Evolution, Learning, and Information*. Oxford, Oxford University Press, 208 p.
<https://doi.org/10.1093/acprof:oso/9780199580828.001.0001>
- STOUT, G.F. 1931. *Mind & Matter*. Cambridge, Cambridge University Press, 342 p.
- VARELA, F.; THOMPSON, E.; ROSCH, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MIT Press, 308 p.
- VARELA, F. 1996. Neurophenomenology: A Methodological Remedy for the Hard Problem. *Journal of Consciousness Studies*, **3**(3):330-349.

Submitted on February 20, 2018

Accepted on May 04, 2018