Psychosemantics and Representationalism[*]

Alexey Chernyak

Faculty of Humanities and Social Sciences
Department of Social Philosophy
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya Street, Moscow, Russian Federation 117198
E-mail: abishot2100@yandex.ru

Abstract—**Psychosemantics is a research program aimed to naturalize meaning, demonstrate how it emerges from natural properties of mental states and processes. The present article explores what seems to be an essence of psychosemantic theories of meaning in order to bring to light their most general problem. It is stated that psychosemantics may give some useful patterns of meaning, but they may become useful as explanations of what mental contents literally mean to a receiving or consuming system, and what the later literally understands from them, only if we abandon the representationalist account of mental contents.**

Keywords—psychosemantics; meaning; representation; intentionality; atomism; holism; proper function

I. INTRODUCTION

Psychosemantics seeks to explain, in the first place, meanings of mental states understood as mental representations. The main idea was borrowed from F. Brentano according to whom meanings of mental states are what they are about or directed at; in other words, their meanings are their intentional meanings [1]. The crucial question then is: how mental states acquire their intentional meanings? [2. P. 259].

In psychosemantics the term 'mental representation' is used to coin what is supposed to be the main features of internal psychological bearers of meaning, viz.: such bearer must be mental and have content capable of representing the intentional object of the state. This presupposes that mental representations must play in thought the role analogous to that of a symbol. And this presumption is often combined with an empiricist claim that some mental contents must directly represent external stimuli (things or events belonging to the subject's environment) which cause their bearers (certain tokens of mental states); they are thus considered as internal bearers of stimulus meanings.[1] In effect mental representations are seen as sorts of symbols which stand for objects and properties of an outside world in thoughts, beliefs and other attitudes.[2]

But how can mental contents represent real things and properties? In particular, how can sensations have external objects as their referents? Psychosemantist normally does not hesitate to say, for example, that the frog's retina 'tells' the brain "there is a bug", and that the brain 'understands' what is thus 'said' [4]. If this metaphor is sound then it has to be explained how exactly retina can inform the brain about what's going on outside it, in the real world itself, and how can brain assign precisely this meaning to the information provided by retina [2. P. 260].

II. ATOMISM VS. HOLISM

One of the most important oppositions in psychosemantics is that between atomism and holism. According to atomism a representation means certain thing because it is caused by the existence of that thing in the outside world [2. P. 261]. Each such representation has a single meaning independently of all others. Thus from the atomistic point of view actual bugs activate frogs' bug-detecting mechanisms and ensure that they communicate information about bugs, or, in other words, that their results mean 'bug' or perhaps 'there is a bug here now' by causing them. This account encounters a number of famous problems though. First, it is unable to explain misrepresentations, i.e. situations where mental content inform the subject about something different from what actually caused it. For example, any small moving black spot of a decent size can activate a frog's 'bug detector' in the same way as bugs normally do. Since the sensations which such activity produce bear some information about what is going on out there they seem to be meaning the same as sensations rightly representing the external world mean in spite of being misrepresentative. But if actual causal chains from an external referent to the interpreted sensational content must determine what this content is about misrepresentations cannot indicate bugs.[3]

---

[1] See [3].

[2] Abstractions and fictions can also be that which mental representation may be about, though.

[3] Thus, according to J. Fodor, there is an asymmetric dependence of meanings of misrepresentation on meanings of proper representations: a representation of a horse may be caused by things other than horses because representations of horses are normally reliably caused by horses [5]. But there is no empirical evidence that this dependence dominates.

Another problem concerns synonyms: they by definition have the same referent; but mental representations may be not known to be synonymous by their possessor. Therefore they can play different roles in the subjective cognitive system in spite of the fact that they refer to the same object [2. P. 263]. Still according to atomism representations with the same external referent should be ascribed the same meaning [5. P. 74]. But then the theory seems to be unable to solve the famous Frege's puzzle [6]. Besides, atomism is unable to account for the existence of abstract concepts (including truth) and thoughts about non-existent entities and states of affairs, counterfactuals in particular [2. P. 263]. For they just cannot be determined by causal links to anything existing in the outside world.

Further, atomism cannot explain why a normal bug detector should demonstrate such perfect selectivity as to be responding to bugs only (at least ideally). If bug detectors respond to whatever looks like a bug, they then are rather detectors of any bug-like stuff, and may be only roughly selective relative to bugs [2. P. 263-264]. But in this case meanings of mental representations have to be approximate too, perhaps disjunctive. Yet it may be insisted that in order to, e.g., make a frog catch primarily bugs, which is crucial for their survival, meanings of mental representations are better be as exact as possible.

And of course there is an additional problem generated by the indeterminacy of causation. Since any mental activity is produced by a chain of causes[4], what reason do we have to pick any single event within the chain (even the first one) as what is (exclusively) meant by the representation?[5]

Holism, contra atomism, assumes that representations cannot have meanings in isolation; they represent a system of inter-related concepts. All representations are connected in a semantic network where elements may be related by different types of connections [2. P. 265-266]. This account can successfully deal with representations of abstract concepts, fictions, generalizations, counterfactuals and so on, since these are according to it of the same kind as representations of concrete objects and events.

But it has its own drawbacks. Thus holistically described, no sign or stimulus has just one meaning; it always means many different interrelated things, and which set of them will be activated depends on the context. Moreover, according to W.V.O. Quine [8], in order to understand the meaning of a single sign or stimulus one needs to know the entire language (or conceptual scheme). Furthermore, since a mental representation only has meaning within a semantic network of an individual mind, and different people have different minds, it has to be explained how two or more people can hold the same belief or another attitude [5. P. 55-57]. It should be then either agreed that they don't, and what we consider to be a communication of one's thoughts to other people and understanding of what other people mean are anything but not what we think they are, or accepted that

there is something independent on those different minds which is common to certain beliefs held by different people (something like broad or wide content).[6]

And finally a semantic network operating with arbitrary symbols may be completely coherent and yet consistent with the existence of a variety of different worlds contradicting between themselves. In order to include representations in a stronger sense, those of something external to the system itself, it has to be somehow grounded or anchored in the external world.[7]

III. TELEOSEMANTICS AND OTHER IMPROVEMENTS

Some thinkers propose a hybrid theory combining aspects of atomism and holism which could solve the problems of both [11]. This postulates an existence of two types of representation: perceptual representations which acquire their meanings due to causal links with an outside world (are directly caused by their external referents) and conceptual representations which meanings are determined by the functional relationships within the system of representations. While solving some problems of the previous accounts this theory still does not prevent as indeterminacy of causally determined meanings so instability of holistically determined ones; and it does not solve misrepresentation problem, neither makes it communication of holistically formed concepts from person to person better explainable [2. P. 266-269].

Another popular solution is *teleosemantics* which combines psychosemantics with teleological functionalism [12], [13], [14]. The general idea is that representations and mechanisms that create them exist because only creatures that possessed them survived the natural selection; and that happened because they served a useful role for those creatures in the past. That past role determines then the proper or ideal function of a representation, i.e. what it is for or what it is supposed to do. The misrepresentation then is explained as a failure of a representation to do what it is supposed to do: a bug-detector activated by a black spot misrepresents a spot as a bug, but it is still a bug-detector because it was evolutionarily selected to detect bugs [13. P. 269].

But in a world where, all other things being equal, shmugs looking and behaving just like bugs (for frogs) exist instead of bugs, the mechanisms internally indistinguishable from actual bug-detectors might be used by frogs for the purpose of nutrition and proved to be as functional for them in that respect as actual bug-detectors are for actual frogs. But there still is a strong reason to count such mechanisms as shmug-detectors, since they were evolutionarily selected for detecting schmugs in that world. Then, why shouldn't we consider rather bug-or-shmug-or-[whatever else might work as good as actual bug-detectors for actual frogs]-detecting as the proper function of actual mechanisms which frogs use for bug-detection? But if so then what guarantees that even if

---

[4] Which, if the determinism is correct, may be traced back to the Big Boom.

[5] The problem was formulated as early as in [7].

[6] See [9].

[7] This is known as the grounding problem; see [10].

these mechanisms have a proper function it makes their results representing and meaning bugs?

Moreover, it seems that instead of explaining away the indeterminacy of meaning teleosemantics generates more of it. Even a small change in the environment can give birth to a completely new uses for a mechanism, and hence to a completely new representative function. Then the representation may be understood either in terms of its original or in terms of its more current function [2. P. 272], [15. P. 401-412].

Teleosemantics may still be said to be preferred because it has certain advantages in comparison with its rivals. Thus it is claimed to explain how creatures often possess multiple mechanisms with similar functions: this happens because it takes some time to adapt to a new environment, and in situations when the environment itself keeps changing, the biology may never catch it up. Then different mechanisms with the same function may evolve. Then while "an imperfect mechanism individually conveys imprecise meaning, multiple overlapping representations can give not just increased precision but also flexibility of choice, depending on which representation (or which combination) is used by the other representations active in the occurrent context" [2. P. 271].

But how this higher degree of performance and hence representational precision may be achieved? If by arbitrary or context-sensitive variation of functional alternatives, then this can have the supposed effect only if 1) the choice-making mechanism (responding for switching between alternatives) is itself close to perfect, i.e. chooses only or mostly what is objectively better to do in the current circumstances, and 2) individual mechanisms between which the subject chooses (or the creature of a more primitive design switches) are real functional alternatives, i.e. have overlapping functions. In reality, though, both conditions are far from being normally satisfied.

And what in the described scenario of evolution makes different tools in a kit having overlapping functions? According to the model each tool was evolving relative to certain evolutionary situation or state of the world but failed to become fully useful,[8] since the state of the world had changed; if proper function is what relates an item to certain set of circumstances or state of a world to which the item was 'designed' to adapt its bearers then the mechanisms which we consider as being functional alternatives, but which were not in fact fully evolved, cannot literally have the same proper function. The situation of real overlapping of functions presupposes that both tools in a kit with overlapping functions are successful selections relative to current circumstances (state of the world) which is not so when the environmental changes are never caught up by biology. In this later case the success of the one tool rather would make another one, which was a good selection relative to some past state of the world (which have changed

radically enough), obsolete. The obsolete mechanism may still be available for some time, but provides no functional alternative to mechanisms which substituted it; then by using such an atavistic device from time to time a person or a creature would rather make the resulting representation fuzzier than improve its precision.[9] Suppose that the threat by showing one's teeth which was supposedly a good choice at some remote stage of human evolution is still used from time to time. Now, when a number of more sophisticated tools of expressing emotions was elaborated, the showing one's teeth to the opponent would rather confuse the addressee than make the message clear; for during the time which has passed since the showing one's teeth had been an action with clear meaning this gesture acquired a number of meanings which associate it with different other actions (intentional or not). So by using such mechanisms one seems to be creating just more semantic indeterminacy.

Besides, teleosemantics doesn't prove that what a mechanism with a proper function (imperfect or not) provides is meaning (imprecise or precise), and not just the activation of certain behavior. What's going on when evolution makes certain mechanisms selected to the benefit of their bearers relative to current circumstances or environment may be described without any reference to semantics at all. A mechanism or a variation of them is selected due to their functional properties relative to a number of certain biological systems such as nutrition, orientation, movement etc. They may be said to be literally selected or designed just to satisfy certain biological needs, not to represent that or this sort of objects; then they may be said having nothing to do at all with semantic meanings.[10]

There are theories, aimed to cope primarily with the problem of indeterminacy of meaning, according to which either evolution or ontogeny may determine a mechanism's proper function; but what frogs' bug-detector detects depends on what exist in the current environment of its bearer (whether there are bugs).[11] But why should we think so? This is left unclear.

Rose adds to this theory the idea of a nested hierarchy of systems which nature spontaneously forms, where a system is defined as a collection of interacting parts (each of which is also a system) generating at least one new emergent property not possessed by any part in isolation. For each system, units with which it immediately interacts constitute a certain level, and systems interacting at one level generate a higher level system. Every event then is said to exist at multiple levels [2. P. 273, 275]. Thus meaning as understood by this theory is a multi-leveled structure.

---

[8] I.e. such that it could be checked out whether it serves to the survival of its bearers in given circumstances (those to which it had to adapt those creatures).

---

[9] Besides, the problem of fuzzy representations seems to embrace perfect mechanisms as well as imperfect.

[10] Of course, such evolutionarily successful mechanisms by activating a behavior which (in ideal situation) provides all and only what is needed in order to effectively deal with certain sorts of things and/or circumstances, may be said to be making those things meant by the representations generated by those mechanisms in *some* sense of meaning: i.e. they are meant as proper objects of certain biological needs.

[11] See, e.g., [16], [12. P. 272].

But why should new meanings emerge together with any new level of organization of nature? Indeed, every new system creates, at least in theory, a new possible way of thinking about things and events etc. involved in it; but yet thinking as well as meanings-assignments may, and seems to be only available at a particular level of organization of natural processes. Then we must have an independent reason to assign to a frog that kind of thinking which we suppose making our bug-concepts meaning bugs (if there are bugs out there). But it does not seem that we have any such reason.

And why should we think that events exist at different levels anyway? This clearly depends on how a being an event is defined. Ross uses the following definition: "An event is a relatively stable state separated from the preceding and following events at its level by relatively rapid transitions" [17]. But what "relatively stable" as applied to a state should mean? What are minimal conditions of being an event in time? Any state of a system is fixable only if the presumption that nothing have changed during certain period of time is held. What would prevent us from dividing any chosen event onto modules consisting of some temporal part of this event and something which had happened just before this part happened? This will give us a module which is not part of the originally chosen event, but which also could be chosen as an event. Thus the very choice of events is (at least partly) arbitrary, and therefore events are rather constructions created by observers than natural things to be found there in the world and having such and such natural features. And the same seems to be true of the choice of a level of events. We may say that the event of detecting a bug consists from an occurrence of some bug in the vision field of a frog only or from this plus whatever the occurrence of a bug and the transmission of its image to frog's brain consist of, or from all this plus the diachronic constituents of the mentioned constituents, etc. In other words an event may be or be not a multileveled construction, may be or be not constituted by sets of systems depending on how we choose to define it. And still any of these we may count as events supervening on other events but not including them as their parts.

If bug detection includes, say, activating of a neuron A, part of what it in theory means may be said to be an activation of a neuron A. The same may be said about whatever other process are going on at lower levels which bug detecting covers. All these may be then said to give rise to a novel property of a frog: the bug-detecting itself. Then, by assuming that bug-detecting generates a new meaning, it may be said that the current level of organization of processes within frog's body makes bug-detecting firings mean bug-detecting.[12] But how this can provide a frog with the concept of a bug? Nothing in the multilevel theory of naturalized meaning explains this (beyond the explanation already provided by atomism).

What any developmental processes may ensure in the first place is that a tool serves its purpose for a consumer. If two systems - the producing and the consuming - co-evolve there may be a perfect equilibrium of functionality or appropriateness of a tool relative to a context. But it is still unclear how semantic meaning evolves from that. Thus if a frog's bug-detector is supposed to do nothing more than detect bugs then some conceptual mechanism should be considered as a consuming system corresponding to this activity. But if we suppose that bug-detecting should provide food in the first place then a feeding mechanism would rather be corresponding to bug-detecting as its consuming system. The fact that bug-detectors are connected (even if directly) to 'bug catcher' modules may be an evidence that it was developed in order to provide proper inputs for this module [2. P. 279]. But from this it does not follow that the detector's firing should mean 'bug' or 'bug present now' to the bug-catcher. It could effectively supply it with schmugs, be the world slightly different. Why does not it mean rather 'black spot is present now' or 'food is present now'?

The theory proposed that "within the bug detector is a (lower-level) mechanism for detecting small black spots of light on the retina (perhaps, an 'of' centre-surround bipolar cell), and that mechanism's activation means 'dark spot on retina' to its receiving mechanisms…" [2. P. 279]. But this doesn't solve the problem, since what is dark, and what is spot, for instance? How dark and how small an item on a retina should be in order to be a proper input of a certain consuming system? We can imagine the situation when all bugs became slightly bigger and slightly lighter than what were normally identified by a considered subsystem as dark spot on retina. In such situation the subsystem would fail to provide its consuming system with right inputs. But if there are schmugs around which satisfy the condition of causing dark and spotish enough images on frogs' retinas the contribution of this subsystem into the bug-detecting one would still make it useful for providing bug-catching modules with inputs proper for activating them as well as bugs did.

Even if we can say that a subsystem somehow 'means' 'dark spot on retina' to something identified as its consuming system, this wouldn't make the higher-level system 'mean' 'bug' or 'bug is present here now' to its consuming system. Bug-catcher might be evolved in the world dwelled with bugs, and hence being designed to catch bugs; but since it could as well catch schmugs it might be rather defined as bug-or-schmug-catcher or buglike-stuff-catcher, or whatever-causes-dark-spot-on-retina-catcher.

But Ross claims that "meaning can be ascribed to sub-personal activity, which includes neural firing, percepts and thoughts" [2. P. 281]. Of course each of these elements has only partial knowledge and capabilities, compared to a whole person; but it is proposed that the more elements that can work together, "the more profound their collective 'understanding' can become" [2. P. 280].

This could be accepted if we had reasons (and not only the will) to expand the notions of meaning and understanding so as to make it legitimate to say about whatever interaction between systems in which one system provides some information to another that the first literally provides some

---

[12] But mean to whom? To the frog's brain, may be? But is brain the right type of device for meaning assignments? Or to the frog's mind? But do we have any reasons to consider frogs as beings with minds capable of understanding that bug-detecting takes place?

semantic meaning to the second, and the second literally understands what the first thus communicates. But usually such extended uses of "mean" and "understand" look arbitrary. A person whose understanding is supposed to be made more profound as a result of a collective semantic work of his/her subsystems, as a higher-level system, cannot 'read' the information provided precisely for her/his subsystems: only certain subpersonal structures are designed to 'read' it; therefore he/she cannot just infer his/her understanding from what these subsystems 'collectively understand'. A person as a whole is designed (if the term is ever relevant) to deal with certain level of information, not with any information which flows through its internal systems.

IV. CONCLUSION

That psychosemantics needs to be further elaborated in order to explain meanings in their ordinary sense of what is assigned to linguistic expressions and representations in general in communication and thinking is well recognized by the proponents of the account. The problem is not that we cannot understand meanings psychofunctionally or teleofunctionally; we surely can if we like. It is not even the problem that neither theory within the framework can solve the problem of indeterminacy of meaning; for which theory can? The real problem with the account, as it seems to me, is its commitment to representationalism. The very idea that mental semantic contents, especially perceptive, have to represent some states of a world or an environment requires that semantic properties should be somehow determined by those external states (at least be true or false relative to them). But then some mental items have to be related to things in the world themselves so as to provide representations of the desired sort by contributing certain content into mental complexes which they constitute together with other items. In simple words the content generated by a bug-detection should refer to a bug if it has to represent a bug's presence. But no bug-content may be literally determined by real bugs or whatever other purely external staff. No bug can become part of a mental design of a person or a creature, and no mental power can extend the limits of the body of its bearer[13] in order to capture something from an outside world or to stick to it somehow.[14] An image of a bug can be internalized, but it itself must refer to a bug for the one who has it in order to represent a bug by being used in a mental activity.

Certainly psychosemantics may give some useful patterns of meaning, but they may become useful as explanations of what mental contents literally mean to a receiving or consuming system, and what the later literally understands from them, only if we abandon the representationalist account of mental contents. The semantic network does not need being about something external at all. It may be practically useful, providing for a creature or a population a way of life by which it manages to survive, adapt, proliferate etc. in a given environment, stable or changing, without being strictly representative relative to that environment or to the one which it was evolved in. And different persons or creatures may have similar semantic networks and concepts due to how they are internally designed, not due to there being certain external things interactions with which make those networks similar.

REFERENCES

[1] F.C. Brentano, Psychologie vom empirischen Standpunkt, Band 1. Leipzig: Duncker und Helmblot, 1874.

[2] D. Rose, "Three Dimensions of Psychosemantics", Gestalt Theory Vol. 34, No. 3/4, 2012, 259-286.

[3] J.A Fodor, The Language of Thought, Hassocks: Harvester Press, 1975.

[4] J.Y. Lettvin, H.R. Maturana, W.S. McCulloch, W.H. Pitts, "What the frog's eye tells the frog's brain", Proceedings of the IRE 47, 1940-1951, 1959.

[5] J.A. Fodor, A Theory of Content, and Other Essays, Cambridge, MA: MIT Press, 1990.

[6] G. Frege, "On Sense and Reference", Translations from the Philosophical Writings of Gottlob Frege, ed. P.T. Geach and M. Black, Oxford and New York: Blackwell, 3rd edn, 1980.

[7] H. Reichenbach, Philosophie der Raum-Zeit-Lehre, Berlin: de Gruyter, 1928.

[8] W.V.O. Quine, Word and Object. Cambridge, MA: MIT Press, 1960.

[9] T. Burge, "Individualism and the Mental", in P. French et al. (eds) Midwest Studies in Philosophy, vol. 4, Minneapolis, MN: University of Minnesota Press, 1979.

[10] S. Harnad, "The symbol grounding problem", Physica D 42, 1990, 335-346.

[11] F.I. Dretske, Explaining Behavior. Reasons in a world of causes, Cambridge, MA: MIT Press, 1988.

[12] D.J. Buller, Adapting Minds. Evolutionary psychology and the persistent quest for human nature, Cambridge, MA: MIT Press, 2005.

[13] D. Rose, Consciousness. Philosophical, Psychological and Neural Theories, Oxford: Oxford University Press, 2006.

[14] G. Macdonald, D. Papineau (eds.), Teleosemantics, Oxford: Oxford University Press, 2006.

[15] D.C. Dennett, Darwin's Dangerous Idea. Evolution and the meanings of life, London: Allen Lane, 1995, 401-412.

[16] J. Garson, Selected effects and causal role functions in the brain: the case for an etiological approach to neuroscience, Biology and Philosophy 26, 2011, 547-565.

[17] S.A. Kaufman, The Origins of Order. Self-organization and selection in evolution, New York: Oxford University Press, 1993.

[18] D. Chalmers, E. Clark, "The Extended Mind", *Analysis* 58: 10-23, 1998.

[19] H. Putnam, "The Meaning of Meaning", *Philosophical Papers*, vol. II, Cambridge, and New York: Cambridge University Press, 1975.

[20] T. Burge, "Individualism and the Mental", in P. French et al. (eds.) *Midwest Studies in Philosophy*, vol. 4, Minneapolis, MN: University of Minnesota Press, 1979.

---

[13] But compare the defense of an opposite view in [18].

[14] It is sometimes claimed that meanings are not in the heads, that mental representations have broad contents which are in fact certain states of affairs in an outside world; see [19], [20]. But it is hard to see how having a broad content may be something more than a metaphor; made literal, it would say, as it seems, either that a content must be about certain things or refer to them, or that contents are identical to states of affairs which we ordinarily suppose being represented by them.