

# AN INTERVIEW WITH JOHN SEARLE

APRIL 1997

University of California, Berkeley

**T**raditionally, THE DUALIST attempts to include an interview with a prominent philosopher in every issue. This year, we were most fortunate to have Prof. John Searle of UC Berkeley agree to participate in such an interview. In an attempt to put a novel spin on the interview format, we asked Searle to field questions from the philosophy faculty of Stanford University.

John R. Searle is Mills Professor of the Philosophy of Mind and Language at the University of California at Berkeley. Searle's research interests include the philosophy of language, philosophy of mind, metaphysics and the philosophy of action. He is the author of numerous books and articles, including: *Speech Acts* (1969), *Expression and Meaning* (1979), *Intentionality: An Essay in the Philosophy of Mind* (1983), *Minds, Brains and Science* (1984), *Foundations of Illocutionary Logic* (with Daniel Vanderveken) (1985), *The Rediscovery of Mind* (1992), and *The Construction of Social Reality* (1995).

**SOLOMON FEFERMAN:** In your recent review of David Chalmers' book *The Conscious Mind*, you say that the correct approach to an account of consciousness is "to forget about the obsolete Cartesian categories and keep reminding ourselves that the brain is a biological organ, like any other, and consciousness is as much a biological process as digestion or photosynthesis." Can philosophy, then, hope to have anything significant to add to the study of consciousness beyond what the biologists (neuro-physiologists, etc.) succeed in determining about it?

**JOHN SEARLE:** There are so many philosophical problems surrounding the notion of consciousness that I believe this will continue to be an important topic in philosophy, even after biologists have answered what I regard as the most important question, namely, "How exactly does the brain cause all of our conscious states?" Here are some of the remaining questions. What exactly is the phenomenal and intentional structure of consciousness? What exactly is the relation of consciousness to intentionality? What exactly is the unconscious and what is its relation to consciousness? What exactly is the causal role of consciousness? How exactly do the conscious mechanisms work in relation to other neuro-biological mechanisms? I believe that regardless of the progress we make in neurobiology, there is going to be a huge philosophical agenda that will occupy us for many years to come.

**EDWARD ZALTA:** In your recent work on the nature of consciousness, you seem to commit yourself to the following three propositions:

- (1) Consciousness is a biological, and hence scientifically reducible, phenomena.
- (2) The subjective aspect of mental states is not a scientifically reducible phenomena.
- (3) Consciousness just is the subjective aspect of mental states.

**These propositions, however, constitute an obvious contradiction. Can you please explain why you haven't contradicted yourself?**

The apparent contradiction that Zalta points out is resolved by pointing out an ambiguity in the notion of scientific reduction. I distinguish between eliminative and non-eliminative reductions, and within the category of non-eliminative reductions, I distinguish between causal reductions and ontological reductions. Now, the situation with consciousness is just this: We can do a causal reduction of consciousness but not an ontological reduction. That is just a fancy way of saying that we can give a complete causal account, in principle at least, of how brain processes cause conscious states. But having done so, we have not given an ontological reduction, we have not shown that consciousness is nothing but neuronal processes in the brain. Nor can we do an eliminative reduction. The causal analysis does not show that consciousness does not really exist. So, consciousness, as Zalta is correct in seeing, is different from other biological and physical properties. It is typical of scientific reductions that we can give an ontological reduction following on the causal reduction. Think of heat or solidity, for example. The reason that we cannot give an ontological reduction of consciousness can be stated quite simply: Consciousness has a first-person or subjective ontology. It exists only as experienced by person or subjective ontology. It exists only as experienced by a conscious agent. Because consciousness has a first-person ontology, it cannot be reduced to anything that has a third-person ontology.

That, in a nutshell, is the reason why consciousness, though an ordinary part of our biological lives, cannot be reduced to other biological processes. We are not used to thinking of anything that has a first-person ontology as being an ordinary, natural phenomenon, but that is exactly what consciousness is.

**FRED DRETSKE:** Over a long and distinguished career you have (in my view) gotten many things right. You have also (once again, in my view) got some things wrong. Quite aside from what I think, though, what, in your own estimation, has been your most serious philosophical mistake?

Dretske asks me what is my most serious philosophical mistake. I have made so many philosophical mistakes that I would not be in a position to say which one is the most serious. Fortunately I have had a lot of enthusiastic help from my fellow professionals in pointing out my mistakes.

**DAGFINN FOLLESDAL: (1) How can one know, when one translates French ‘lapin’ into English ‘rabbit’ that this is the right translation and that no other translation is correct? (2) What do you regard as the main advantage(s) of your own theory of metaphor over that of Davidson?**

(1) The problem of understanding another person’s speech in a foreign language is an example of the problem of understanding another person’s intentional behavior in general. Because speech acts are intentional acts, the problems of understanding speech acts is part of the general problem of interpreting behavior.

The problem of understanding what words conventionally mean in a foreign language is a problem of understanding the conventions of that language and how they map onto the conventions of English. There are, of course, certain difficulties that arise in real life; for example, many words of a foreign language do not translate exactly into English. The French word ‘*esprit*’, for example, does not translate exactly as ‘mind’.

Follesdal’s question was how do I know that “this is the right translation and no other translation is correct?” Well, in any given case, one could continue to have skeptical doubts, just as one can have doubts about the interpretation of other people’s behavior in general. But in most real life cases, the doubts are unrealistic. If somebody asks me how I know that ‘*lapin*’ means ‘rabbit’, I can only say that I have been speaking French for over forty years, and I know French well enough to know that this is what I mean when I use this word, and that I share this meaning with other French speakers.

I believe that there only appears to be a philosophical problem about translation because we are still suffering the bad effects of old time behaviorism and verificationism. The problem of ascertaining the meaning of other people’s words is just a special case of the other minds problem. If, however, you are tempted to think that there can be no facts about meaning which are not facts about overt behavior, then it is possible to produce a *reductio ad absurdum* of your behaviorism, because you find that you cannot distinguish on behavior alone between meaning ‘rabbit’ and meaning, for example, ‘undetached parts of a rabbit’. Quine produced a *reductio ad absurdum* of his behaviorism many years ago with his argument for the “indeterminacy of translation.” So, one has a choice. As Gil Harman says, “One man’s modus ponens is another’s modus tollens.”

er man's modus tollens." In this case you can, with Quine, hang on to behaviorism come what may, and insist that there is no fact of the matter about whether when I say 'rabbit' I mean rabbit or undetached parts of a rabbit. Or you can, with me, acknowledge what I take to be the obvious facts that there is a clear distinction for me, between my meaning rabbit, and my meaning undetached parts of a rabbit, and that any behaviorist account of meaning that cannot acknowledge those facts is therefore false.

(2) If I understand him correctly, Davidson does not, in my sense, have a "theory" of metaphor. Davidson agrees with me that the words in a metaphorical utterance have their ordinary literal meanings. There is no change of word or sentence meaning in a metaphorical utterance; for if there were, the word would have ceased to be used metaphorically and would simply have acquired a new literal meaning. It would, in short, be a dead metaphor. Where I believe my theory has the advantage over Davidson's is that I try to explain the fact that metaphorical utterances have a propositional content that is not expressed by literal meaning. The distinction between the speaker's utterance meaning and the literal sentence meaning applies to metaphorical utterances. The metaphorical part is in the speaker's utterance meaning. Thus, if I say, "Sam is a pig," I may assert a true proposition, even though it is not literally true that Sam is a pig. That is, I may use a sentence to convey a metaphorical speaker's utterance meaning when that sentence literally means something different from what I am conveying.

**JOHN PERRY:** Consider a property like *believing that Searle is a philosopher*. Suppose we agree that this is a real property that a person might have, and that it plays a role in the causal order of things and is relevant to explanations, and that it is a perfectly respectable, naturalistic, scientific property that does not need to be identified with or reduced to any other physical property. These are all parts of your view that appeal to me.

Here is a question about the structure of that property. Suppose we agree further that the property in question has in some sense a proposition as a constituent. There seem to be (at least) two alternatives, which I'll call *a* and *b* for lack of better terms.

*a* is a conception that Frege articulates in "The Thought." Propositional attitudes involve our grasping an abstract object. Belief is basically a form of grasping, and so a relationship between the believer and the believed proposition.

*b* is a conception from "two-tiered" accounts. Propositions are not *relata* of mental relations, but abstract objects used to classify mental states and events (as well as other sorts of information-rele-

vant states and events, such as assertions). So the property *believing Searle is a philosopher* is a bit like the property *weighing 200 pounds*. One can think of weighing as a relation between a thing and an abstract object, 200 pounds, but that would be somewhat misleading. It is rather that the person has a complex property that one abstract object that is part of a system of mathematically related abstract objects can usefully classify. Mathematical relations between the abstract objects map onto physical/natural relationships between the objects.

I like *b*. So my questions are:

- 1) Do you hold *a* or *b*?
- 2) Could someone hold your views, as described in the first paragraph, and still hold *b*?
- 3) Do you see any interest or value in *b* whatsoever?

In answering a question like Perry's, my approach is always to ask, "What facts in the real world correspond to the claims we are making?" What fact in the world makes it true to say, "Perry believes that Searle is a philosopher"? In the first instance, it seems to me that the fact that corresponds to this claim is that Perry is in a certain intentional state, the state of believing that Searle is a philosopher. I say 'state', because the grammar of 'believe' is that of a state verb and not a process verb. For example, in English 'believe' does not take the present continuous tense. As an answer to the question, "What are you now doing?" we cannot say, "I am now believing that Searle is a philosopher." But what is this intentional state? Well, under investigation it turns out to have both logical or representational properties, and also psychological and neuro-biological properties. The former properties are realized in the latter. The logical properties are describable in terms of propositional content, direction of fit, conditions of satisfaction, intentional mode, etc. I have written a book about these, and it is all a rather long story that I will not repeat now. The point for the present discussion is that all these logical properties are realized in the plumbing. This whole intentional package is realized in his brain, and as so realized, it has certain causal properties. It is capable of being in the form of a conscious thought, it will function causally in producing certain sorts of behavior, etc.

Now, with all these facts in mind, what about his options *a* and *b*? I do not accept either of these. I can see how someone suitably confused might be tempted to *a*, but I don't understand *b* at all. The analogy with weight is of no help, since my weight is a gravitational relation between my mass and the center of the earth, not between me and mathematical objects. I don't for the life of me see what could possibly motivate *b*. So, my own approach, to repeat, is to try to stay as close to the obvious facts as I possibly can.

**JULIUS MORAVCSIK:** (1) Causal, externalist theories of knowledge or language interpretation can be successful only if nominalism or finitism is true. But some of the best minds of this century (Hilbert, Quine, Lesniewsky, etc.) have tried to make those theories go, and failed. So why do analytic philosophers push “naturalistic” theories of reference, belief, knowledge, etc. when as things stand, they cannot explain even something as simple—within their theories—as “Jones believes (knows, understands) that  $2 + 2 = 4$ ”?

(2) The dualism of mind and matter has a sharp delineation only in Cartesian and post-Cartesian philosophy. In *that* sense of “matter” (extended bits of stuff), however, not even physics is materialistic. So why do these philosophers try with great effort to climb aboard the sinking ship of materialist reductionism or physicalism? Quine said, “Philosophers tend to confuse the compartments of the universe with the departments of the university.” Are not physicalists guilty of this?

(3) Analytic philosophers make the extremely pretentious claim that their analytic definitions are, when true, true across “all possible worlds.” But if human understanding is partly restrained to historical intellectual contexts, how could the human mind at any temporal stage survey “all logically possible worlds”? Your own recent work seems to be moving in the direction of taking account of social-intellectual context in interpreting and constructing theories. So where do you stand on this issue? Should we not just talk about a contextually restricted subset of all possible worlds?

All three of Moravcsik’s questions are about the current state of philosophy and the motivations of other philosophers. Since I don’t hold the views that concern Moravcsik, I really am not in a very strong position to speculate about what motivates philosophers who do. But I will try as best as I can.

(1) Why do analytic philosophers push naturalistic theories of reference, belief, knowledge, etc.? I think the answer is that at some point they want a coherent and unified account of the whole universe. They generally accept some sort of “materialism,” and they would like to give an account of the mind, belief, knowledge, etc., which is consistent with their materialism. They want to treat these things as part of nature, hence as “naturalistic.” I think, in a way, the problem is both easier and harder than they realize. It’s harder because given the categories they use, they will not succeed in giving a naturalistic theory of these phenomena. But the problem is also easier, because the simplest way to give a naturalistic theory of, for example, intentionality, is to recognize that intentionality is a biological part of nature. It already is naturalized.

The problem then is to give a logical analysis of its structure, as well as a psychological and neuro-biological account of how it works causally. Any such account is naturalism enough.

(2) The same considerations apply to physicalism. If you believe that mental states are not part of the ordinary material world, then you have a problem. It's the problem of dualism. So, the temptation of philosophers is to reduce consciousness and other mental phenomena to physical phenomena. All of these efforts fail for reasons that I have tried to explain in enormous detail.

(3) I am puzzled by Moravcsik's third question. When analytic philosophers talk about all possible worlds, or all logically possible worlds, he makes it sound as if they must have a huge job of counting. But of course, possible worlds are not things that they hope to count. Rather the possibility in question is logical possibility, and if we can give an account of logical possibility, we don't need to worry about the number of possible worlds. The metaphor of possible worlds is a picturesque way of explaining certain notions in semantics, notions such as what might have been the case, what must be the case, etc. The notion of possible worlds raises no special epistemic problems about knowing about "all possible worlds."

Berkeley, CA  
University of California