

PATRICIA SMITH CHURCHLAND

FODOR ON LANGUAGE LEARNING*

At least since Wittgenstein aired the matter in the *Philosophical Investigations*, the Augustinian theory of language learning has generally been regarded as a quaint old relic, as naive and uncontroversially wrong. It has been virtually a commonplace in contemporary discussions of language that however a language is learned by the aspiring child, it is evidently not through the auspices of an unlearned language over which he innately has command and by means of which he conducts his cognitive ruminations. Surprising it is then, that this very theory has been rescued from the curio shop by Fodor in his latest book, *The Language of Thought* (1975). Surprising also is the strength of the defense Fodor discovers can be marshalled on behalf of the Augustinian view. He argues that its status as a mere curio is undeserved, and that not the least of its virtues is that it is the only decent theory of language learning in contention. The book will be widely read, indeed it should be widely read, by philosophers and psychologists alike, and it contains the potential to influence greatly the direction and development of psycholinguistic research. I hope that its potential to so influence will not be realized however, for, as I shall argue anon, it is thoroughly misconceived.

The undoing of the theory, at least in the form Fodor presents it, is not so much that innate wherewithal is posited; it is rather the consummate richness and fixity attributed to the innate wherewithal. According to Fodor, the innate structure needed for learning a language is an innate language, and this endowed accoutrement is no pale prototype of the language *vocalis* that the child will acquire, nor indeed is it a mere germ which grows and develops to reflect the sort of *Weltanschauung* embodied in the particular language of the child's milieu. Mentalese, Fodor argues, is as rich and powerful, as complex and complete, as any language, be it English or Urdu, the child comes to learn. Fodor is forthright in putting the point:

What, then, *is* being denied? Roughly, that one can learn a language whose expressive power is greater than that of a language that one already knows. Less roughly, that one can learn a language whose predicates express extensions not expressible by those of a previously available representational system. (86)

Synthese 38 (1978) 149–159. All Rights Reserved.

Copyright © 1978 by D. Reidel Publishing Company, Dordrecht, Holland.

What the child cannot do, in short, is use the fragment of the language that he knows to increase the expressive power of the concepts at his disposal. (84)

We must now add, that the same models imply that that language [Mentalese] is extremely rich (i.e., that it is capable of expressing any concept that the organism can learn or entertain) and that its representational power is, to all intents and purposes, innately determined. (97)

Just how Fodor comes to such conclusions, which he acknowledges 'seem scandalous' (82) is worth reviewing. Initially, Fodor argues that in order for the pre-verbal child to learn the predicates of the mother tongue, he must entertain hypotheses, test them against the evidence, revise them etc. Such cognitive enterprise requires, the argument continues, a medium of representation in which the child carries out the hypothesizing, testing, and revising. It requires then a system of representations which the child may exploit in his linguistic endeavours; indeed which he must exploit if he endeavours at all in matters linguistic. Now to this point in the argument, things do not look too bad, inasmuch as it does seem reasonable to regard the child as doing, amongst other cognitive things, something like hypothesizing and testing, and on the other hand, the stimulus-response model just does not seem to go very far in explaining the child's linguistic accomplishments. However, Fodor believes that the stimulus-response model does not go anywhere in so explaining, and moreover that it is not in any attenuated or rarefied sense that the preverbal child hypothesizes; the child literally, in the full-blooded sense, frames hypotheses and tests them.

Having thus construed the cognitive activities of the child, Fodor is moved to try to give an account of what the hypotheses framed by the language learner look like. His conclusion is that they take the form of biconditionals matching Mentalese predicates with predicates of the spoken language. An example of such an hypothesis is the following:

- (1) 'Bear(x)' is true (in English) iff $G(x)$

where 'G' is a Mentalese predicate. The child has mastered 'bear' when he has adopted the correct truth rule; that is, when the extension of 'bear' and the extension of 'G' are the same.¹ Clearly, if the child is to learn the meaning of 'bear' via coming to adopt (1), he must already understand 'G' (80); 'G' must be a predicate in the repertoire of his Mentalese (82). *The sum and substance of language learning thus turns out to be translation.* There is no such thing as

concept learning (96); the child learns English by coming to translate from Mentalese to English, and any English predicate he can learn has a Mentalese Doppelgänger. The only exception here concerns English predicates which are explicitly definable by means of a truth function of other English predicates. Fodor's example here is 'airplane' which he says is definable as 'flying machine'. But learning 'airplane' would not constitute learning a brand new concept since Mentalese must be stocked with the counterparts of 'flying' and 'machine'. Given that language learning (or, more properly, language acquisition) is conducted solely and entirely by means of hypotheses such as (1) (save for the not very interesting class of predicates defined explicitly in terms of other spoken language predicates) the posited Mentalese must be a consummate, fleshed out, full-blooded, but withal unlearned, system of representations.

In certain respects Fodor's theory resembles the homunculus theory of human reproduction. Faced with the puzzle of how a complete human being could come into existence inside the womb, certain philosophers concluded: it does not. The complete human being, so the story went, already exists. It exists within the loins of the male and it is merely transferred to the female enlarging chamber (the womb) wherein it proceeds to grow, *in just the ways* that the born child grows. Similarly, faced with the well-nigh intractable puzzle of how a child comes to learn a language, Fodor counsels: it does not. A complete language is known innately, and the child merely formulates translational hypotheses, *in just the way* his accomplished seniors formulate translational hypotheses between English and Urdu. Now as everyone knows, the homunculus theory was not very successful, and part of what was amiss in the attempt was a certain narrowness of vision or imagination. True enough, nothing was known of chromosomes and their marvellous replication, nor of the exquisite mysteries of cell differentiation, but the narrowness consisted in slavishly modelling growth *in utero* on growth *ex utero*, when *prima facie*, whatever was going on *in utero* had to be radically different. Well it may be wondered whether the model of translation as poorly fits the learning of the mother tongue.

What is so implausible about Mentalese is that it is surpassingly rich, and that it has this richness independently of any learning processes. Fodor does suggest that maturational factors likely have some role in the development of a full-flowered Mentalese, but whatever that role is, it allegedly does not

come to learning (94, 95). The richness together with the innateness constitute a rigidity, a resistance to modification, an imperviousness to learning from one's mistakes that is wholly uncharacteristic of the language of communication. It is not just that the hypothesized inner system of representations should be supposed so different from the observed verbal one. It is that the *language of communication*, used to represent how the world is, should be so sensitive to change, whilst the *language of thought* should be so stable and enduring, and it too be used to represent how the world is.

If Fodor's theory is kindred in spirit to the homunculus hypothesis, so also is it kindred to logical atomism. The predicative elements of Mentalese are the semantic atoms, and truth functional logic is the bountiful machine which produces, when necessary, combinations of the elementary particles. On the question of how we come to use these elements appropriately, Fodor's answer is that we are just built to use the predicates of Mentalese correctly. (66) Moreover, as noted earlier, Fodor insists that we do not augment the repertoire of Mentalese by virtue of learning or any other cognitive process, for we do not learn any concepts. Now the representational system as a whole presumably does eke out the fundamental Weltanschauung of each of us. How we represent the world, what in general we believe to be in the world, what sorts of properties we believe things to have, is exactly proscribed by the innate representational system. Change in the representational system does not consist, so it seems, in the emergence of new concepts and the attendant re-positioning of concepts in the representational structure, but only in new combinations of old concepts. Ostensibly new concepts are therefore always *reducible* to congeries of old concepts; ostensibly new theories about the nature of the world are essentially verbal filigree worked on the basic and abiding theoretical structure. (This thesis gives a new squint to the old saw that there is nothing new under the sun.) Apparently we may tinker and toy with the semantic molecules, combining and decomposing as we will, but we cannot aspire to foment conceptual revolution. After all, we do not learn any concepts, and those we have got, we are simply built to use.

Still, for any Weltanschauung, for any structure allegedly exploitable to represent how the world is, it is pertinent to inquire whether it might not be in certain ways inaccurate, or unfaithful to the facts – whether how it *represents* the world might not be at odds with how the world *actually is*. Of course if a species survives the ravages of natural selection, this does not show

that its representational scheme is accurate, but only that it is good enough for survival. A frog fails to represent a dead fly as a fly (a motionless morsel as an edible morsel), and while a particular frog may starve in the midst of plenty of dead flies, the bountiful supply of flying flies keeps the species going. So far as I can determine, Fodor is unbothered by the matter of accuracy raised in regard to Mentalese, nor does he appear to be impressed with the view that development in science has wrought profound changes in certain of our aboriginal ways of representing the world, and that this has come about precisely because it seemed that these ways of representing the world were inadequate. For all I know, we may be innately disposed to represent space Euclideanly, or to represent the sky as a vault, or to represent dolphins and whales as fish, and so on. But even if we are, I think it is fair to say that we have shown some alacrity and aptitude for reform. Not surprisingly, I suppose, Fodor is silent on the matter of how his theory is to deal with the wealth of theoretical concepts bequeathed us from the sciences, such as 'electron', 'wavicle', 'force field', 'valence' etc. In fact, he has a non-trivial problem with them, and the problem, given his thesis, is this unappetizing dilemma: either the (overtly) theoretical concepts of science are innate, or on the other hand, they are reducible to truth functional combinations of innate concepts. Taking the first alternative, it will have to be admitted that the ancient Greeks, the Huns, and perhaps even Cro-Magnon man, possessed a Mentalese graced with such concepts as 'electron' and 'valence', and the wonder is that the relevant theory was so long in articulation. On this alternative, we attribute to Mother Nature a thoroughly uncanny prescience in outfitting our representational systems, for it appears that the science in our genes scoops, and scoops systematically, the science in our libraries.

The second alternative fares no better, for there is not the slightest reason to believe that theoretical predicates are reducible to elementary predicates, should there be such. Quite the reverse; there is good reason to believe that theoretical predicates are not so reducible. If we suppose that the evolution of language, properly viewed, is nothing but the accretion of neologism via explicit definition in terms of already existing concepts, we contradict everything we know about the theoretical vocabulary of the sciences. Explicit definition may do well enough to explain the conceptual and linguistic niche occupied by 'pediatrician', but it will not do at all for 'electron' and 'valence'.

These and other items from the vocabulary of science are *cluster concepts*², and their meaning is a function of the set of general propositions in which they figure. They do not, then, enjoy eliminative definition at all, but rather are *implicitly defined* by a network of *non-analytic* propositions, which jointly constitute a theory or some part thereof. Since discoveries concerning theoretical entities and properties are made from time to time, and theories are modified and revised, the network in question changes in minor, and occasionally even in transmogrifying, ways. Nevertheless, as Quine and others have long urged, it is vain to try to separate what we mean by 'Ø' and what we take to be the factual truth about Ø's or to try to distinguish between changes in the meaning of 'Ø' and changes in our beliefs about the sort of thing Ø's are. Now what the second alternative enjoins us to accept is starkly incompatible with these views, for what it amounts to is primitive reductionism buttressed by a coarse sort of analytic-synthetic distinction. The dearth of convincing argument *for* primitive reductionism and the abundance of convincing argument *against* primitive reductionism fairly scotches the second alternative.

With both alternatives going to the wall, it follows that theoretical expressions are neither innate, nor are they reducible to truth functions of innate expressions. The conclusion forced on Fodor is that his theory is incapable of allowing that anyone comes by an understanding of science, let alone that anyone can pen new theory in science.

What has to give way in Fodor's theory is the consummate richness and fixity of Mentalese. A place in a theory of language learning must be found for evolution and development, for invention and iconoclasm, in short, for *bona fide* learning. Consequently, what goes too is the story on the form and nature of the cognitive activities of the learner. I expect Fodor may decry this recommendation as horses' medicine, consisting essentially in putting scare quotes around 'hypothesizes' and therewith snuffing the life out of the theory without so much as a hunch to replace it with. Perhaps it does come to that. Yet why resolutely abide with a theory should it become evident that its promise is at the vanishing point?

Fodor seems to make much of his claim that the Augustinian theory stands alone as a decent theory of language learning. Exactly what methodological point he is making, however, is a bit dark. If the point is that as scientists and philosophers we should be open to surprising theories and

surprising results, and that we should give the quietus to rigid pride and unruly prejudice, then the point is well taken. If he is asking us to consider seriously his theory of language learning *because* it is plausible and constitutes a promising framework for further research, again, the methodological point is well taken. However, if he is asking that we take it seriously and try to conduct further research within its framework despite its shortcomings *because* it is the only theory we have got, the methodological justification is wanting. Evidently if his *is* the only theory we have, it is worth determining whether it stands a chance of being right, and if not, how it is wrong. But just as evidently, if it is fraught with refractory woes, the thing to do is to set about inventing competing theories. At one point Fodor says: "Remotely plausible theories are better than no theories at all." (27) Methodologically, this is a remarkably backwards and conservative sort of thing to say. Cleaving stalwartly to an unpromising theory on grounds that it is the solitary contender can surely be utterly unreasonable, inasmuch as the time spent vainly trying to make the thing fly might be better spent devising a new approach. Theoretical vacuums after all, are just the sort of thing to stimulate the faculties. A theory whose demise is delayed by the methodological tenet that remotely plausible theories are better than no theory at all, can simply prolong frustration. Better Feyerabend's counsel (1970) to proliferate competing theories even when a temptingly decent theory occupies the field, let alone when an un compelling one holds sway.

Finally there is something profoundly misleading in Fodor's characterization of the theoretical situation as desolate but for the Augustinian approach. To begin with, a theory of language learning properly belongs within a theory of learning in general. There is no compelling reason to suppose that how a child learns what he learns *before* he learns a language is utterly unrelated to how he learns a language, and to how he learns what he learns after he learns a language. Learning a language is likely an instance of learning – it may of course be a somewhat special instance, but then so may be learning mathematics, learning to play a musical instrument, learning spatial constructions and how to find one's way around one's environs. Moreover, as Fodor would concur, there is no compelling reason to suppose that how the pre-verbal child learns is utterly unrelated to how non-verbal creatures in general learn, and certain it is that non-verbal creatures do learn.

What is important then, in assessing our theoretical prospects vis-a-vis

language learning, is whether there is a programme afoot which holds promise for a theory of learning in general. To this end, consider the following sort of picture: we humans are basically information processors, albeit of a remarkably intricate and splendidly complex variety. *Au fond*, we are epistemic engines³, whose epistemic states are a function of environmental influx and existing internal states. We may regard ourselves as occupying one place in a spectrum of information processing devices which finds a location for porpoises and racoons, and at the nether end, for heliotropes and even voltmeters. Not every creature or artifact on the spectrum merits the honorific 'epistemic engine' in the sense herein used, inasmuch as some things, for example heliotropes and voltmeters, are mere informational sponges, limited sensorily in the intussusception of information and limited also in their behavioural repertoire. Higher on the spectrum, it appears that sensory channels are more diverse and finer-grained, that the information filters are increasingly complex, and the filters themselves are organized into complex systems. As we move into the upper reaches of the spectrum, 'learns' begins to be appropriate in describing the creatures, and with man we encounter a creature learning inveterately and incessantly – a flourishing and rather intriguing epistemic engine.

The temptation, of course, is to take the linguistic utterings and mutterings of *homo sapiens* as the model for all information processing in the cortex (of *homo sapiens* and indeed of lesser beasts), and thus to see each step in information processing as one in which the individual manipulates a proposition in a certain manner (e.g. he hypothesizes, concludes, accepts etc. that *p*). The abominable consequences of Fodor's approach dramatically illustrate the sorrows attendant in succumbing to that temptation. Ironically, an outstanding virtue of Fodor's book is this: it forces us to perceive that a theory of information processing in intelligent animals, rather than building *within* the paradigm of proposition manipulation, will have to look *beyond* that paradigm. P. M. Churchland (1975), in trying to win us away from the view that the model of information processing in general should be fashioned on the model of overt linguistic behaviour, remarks:

... and what I suggest is that an adequate theory of what knowledge is and how it grows will have to be concerned with whatever *is* our fundamental mode of representation and with the parameters of *its* elements. One could of course nurture a hope that the structure, elements, and operations of human language systematically reflect or mirror all of the

theoretically relevant structure, elements, and operations of the brain, but there is no empirical evidence to sustain such a hope, and one would expect on the contrary that linguistic structures/operations reflect brain structures/operations only very grossly, selectively, and superficially.⁴

Hooker (1975) is intent on sowing similar doubts when he says:

Language will surely be seen as a surface abstraction of much richer, more generalized information processes in the cortex, a convenient condensation fed to the tongue and hand for social purposes. The cortical processes themselves, it seems, will be much richer in content than natural language captures. Yet these constitute transformations, and inevitably simplifications, of total stimuli.⁵

It is possible, but by no means certain, that particular accomplishments by certain epistemic engines may be best described as the outcome of the engines' covert manipulation of sentences. It is also possible that particular accomplishments may be best described on the stimulus-response model. What is patently doubtful is that betwixt the two they will lick the platter clean. As to what we can envisage in a theory of learning in general, as to what the natural science of epistemic engines will, in the fullness of time, come to look like, I think one can be rewarded with creditable hunches by examining two sorts of current research: (1) that which undertakes to give information theoretic accounts of the function of relatively simple organisms and of parts of complex organisms,⁶ and (2) that which boldly speculates on the general character of information processing in humans.⁷

Now what I have sketched here is not, needless to say, a *theory* of learning in general, but at the very least it is a guiding conception of how we might usefully approach the problem of concept learning in humans and other creatures (and artifacts). It is an approach which attempts to exploit and enrich the resources of information theory, and it is in fact the guiding conception of a good deal of current research, some results of which are sufficiently impressive to engender a cautious optimism. Before an adequate theory of learning in general will emerge under the aegis of this conception, the imaginative faculties will surely be put to the test devising suitable models of the structure and organization of information processing systems, and a good deal of work may have to be done on the more humble fauna before much can be said about man. Moreover, if the history of science is anything to go by, progress will be non-linear and we can expect many a dead end and dashed hope. Whilst I am not confident that a general theory of learning is

imminent, I am convinced that the information processor approach is pregnant with promise. Anyhow, as Feyerabend remarks (1970), it takes time to build a good theory.

University of Manitoba

NOTES

*I would like to thank Ann Wilbur MacKenzie and Nollaig MacKenzie for criticism and encouragement. I gratefully acknowledge support from the Canada Council, grant no. W750470.

¹ Fodor cites convenience as his reason for claiming that semantic rules match extensionally equivalent predicates rather than intensionally identical predicates. He says that:

... these arguments are neutral vis-a-vis controversy between extensionalist and intensionalist semantics. If you are an extensionalist, then surely you believe that the semantic properties of a predicate determine its extension. If you are an intensionalist, then presumably you believe that the semantic properties of a predicate determine its intension and that intensions determine extensions. Either way, then, you believe what I have wanted you to assume. (60)

Putman (1973) has given us reason to pause in the face of such remarks, but since my objections to Fodor are independent of this issue, I shall not press it here.

² Vide Putnam, 'The Analytic and the Synthetic', *Readings in the Philosophy of Language*, ed. Rosenberg and Travis, Prentice-Hall, 1971.

³ I owe this expression to Paul Churchland who coins it in the paper 'Sentential Epistemologies and the Natural Science of Epistemic Engines', unpublished.

⁴ Churchland, P. M. 'Popper's Philosophy of Science', *Canadian Journal of Philosophy*, 5, No. 1, Sept. 1975. p. 156.

⁵ Hooker, C. A. (1975).

⁶ For example: Gunther S. Stent, 'Cellular Communication', *Scientific American*, September 1972; Eric R. Kandel, 'An Invertebrate System for the Cellular Analysis of Simple Behaviors and Their Modifications', *The Neurosciences Third Study Program*, ed. Schmitt and Worden, MIT, 1974; Allen I. Selverston and Brian Mulloney, 'Synaptic and Structural Analysis of a Small Neural System', *The Neurosciences Third Study Program*, ed. Schmitt and Worden, MIT, 1974; Graham Hoyle, 'Neural Machinery Underlying Behavior in Insects', *The Neurosciences Third Study Program*, ed. Schmitt and Worden, MIT, 1974; Lettvin, Maturana, McCulloch and Pitts, 'What the Frog's Eye Tells the Frog's Brain', *Embodiments of Mind*, W. S. McCulloch, MIT, 1965.

⁷ See for example, William T. Powers, *Behavior: The Control of Perception*, Aldine, Chicago, 1973; Karl H. Pribram, *Languages of the Brain*, Englewood Cliffs, New Jersey, 1971.

BIBLIOGRAPHY

- Churchland, P. M.: 'Popper's Philosophy of Science', *Canadian Journal of Philosophy* 5 (1975).
- Feyerabend, Paul: 'How to be a Good Empiricist – A Plea for Tolerance In Matters Epistemological' in *Readings in the Philosophy of Science*, ed. Baruch A. Brody, Prentice-Hall, 1970
- Fodor, Jerry A.: *The Language of Thought*, Crowell, New York, 1975.
- Hooker, C. A.: 'Philosophy and Meta-Philosophy of Science: Empiricism, Popperianism, and Realism', *Synthese* 32 (1975).
- Putnam, Hilary, 'Meaning and Reference', *The Journal of Philosophy* 70 (1973).