

Linguistics and General Process Learning Theory

Michael Flynn

0. This paper is sort of an extended footnote, with a faint Borgesian flavor¹. What I'm going to do is show how one rather prominent argument in the linguistics literature against one aspect of the research program of behaviorism fails to go through. But I'll also observe that this argument appears to have had no practical effect on linguistic investigations, and that many people seem to assume (tacitly, at least) that this argument fails anyway. So my remarks here don't move the field forward any, but what I hope they do do is help to get us all a bit clearer about where we are.

The argument I'll be examining, given by Noam Chomsky in Reflections on Language (Chomsky 1975), is against a point of view called "general process learning theory", a view that regards one goal of psychological theorizing to be the discovery of laws of learning that hold across species and across domains of acquisition.

Psychological theorizing is by no means a new development on the linguistics scene. It is true, I think, that in most cases the people who have thought about language (including but not limited to people we would call linguists) have done so against the backdrop of a psychological theory that they assumed to be at least on the right track, and the idea was often to see what you could make of language by applying the analytical tools that the given psychological theory made available. Bloomfield (1926) is an example of this. (For some discussion of Bloomfield's views on psychology, see Lyons 1978, chapter 3.) One also in this context thinks of Piaget, Skinner of course, as well as philoso-

¹ Recall that Pierre Menard, author of the Quixote, once wrote a technical article on the possibility of improving the game of chess, eliminating one of the rook's pawns. Menard proposed, recommended, discussed, and finally rejected this innovation. For more on Menard, see Borges 1964.

phers of the 17th and 18th centuries of both the continental Cartesian variety and the so-called British Empiricists. I also think it's true that Chomsky's impact on psychology is somewhat unusual in that the flow of influence is in the other direction; that is, the question is, "If human language is like this, then what must the mind be like?" rather than the other way around.

Be that as it may, Chomsky has been, by far and away, the leading expositor of the implications of linguistics for the study of the structure of the human mind. It goes without saying that the ramifications of this work have been very rich, the pivotal role of linguistics in the "cognitive sciences" being just one indication of its influence.

One of the earliest engagements at discipline boundaries was Chomsky's forceful assault on B.F. Skinner's attempt to extend the domain of behaviorist psychology to human languages.² It's this argument that I want to have another look at. To do this it will be useful to try to isolate several facets of the discussion. I should perhaps reiterate, for the connoisseurs of counterrevolution who I know are out there, that my conclusion will be a modest one. I will not be concluding that after all Skinner was right and Chomsky was wrong. On the contrary, I'm going to assume that this game is over, and has been for quite some time. My goal is to call attention to what I think is an unsolved problem which acquires its interest because it bears on how we regard linguistics as influencing our judgment about the structure of the human mind.

1. The first thing to notice is that some of the objections to behaviorism are (or, at least, were) focused on the validity of

² Two notes on terminology are perhaps in order here. I'm using the word "discipline" in the traditional and rather pedestrian sense that might be used by a dean allocating office space. So I'm putting aside questions like whether or not linguistics is psychology and psychology is biology. Also, when I write "behaviorism", what I have in mind is Skinnerian behaviorism. Nothing much hangs on either of these decisions.

extrapolating to human beings the results obtained with other species. Chomsky is very clear about this in his review of Skinner's Verbal Behavior (Chomsky 1959). For example, he writes (p.554):

The operant is thus defined with respect to a particular experimental procedure. This is perfectly reasonable and has led to many interesting results. It is, however, completely meaningless to speak of extrapolating this concept of operant to ordinary verbal behavior.³

By 1972, the judgment against behaviorism had become much harsher, but the main sticking point was still its application to human beings. In his paper "Psychology and Ideology" Chomsky writes, "In his speculations on human behavior, which are to be clearly distinguished from his experimental investigation of operant conditioning, B.F. Skinner offers a particular version of

³ Of course, 1959 is a long time ago, and I don't mean to be suggesting that Chomsky hasn't modified his views concerning some of the details of that famous review. But it is instructive, I think, and it will be of interest for us later on, to observe that the issue was not the legitimacy of the framework for studying animals but its legitimacy when extended to humans, thus implying a rather large gap between the two groups, a kind of bifurcation thesis, if you will.

For those who like to keep track of such things, it may be of some historical interest to observe that in 1959, Chomsky apparently didn't believe that the application of operant conditioning to some aspects of human behavior was inappropriate. In a footnote (p.553, fn.7) he writes:

In fairness, it must be mentioned that there are certain nontrivial applications of operant conditioning to the control of human behavior. A wide variety of experiments have shown that the number of plural nouns (for example) produced by a subject will increase if the experimenter says "right" or "good" when one is produced ... It is of some interest that the subject is usually unaware of the process. Just what insight this gives into normal human behavior is not obvious. Nevertheless, it is an example of positive and not totally unexpected results using the Skinnerian paradigm.

the theory of human malleability ... His speculations are devoid of scientific content and do not even hint at general outlines of a possible science of human behavior" (Chomsky 1972a:12).

In the meantime, though, behaviorism was being severely criticized from other quarters, raising problems for the framework as a whole by raising doubts about some of its fundamental assumptions. It will be useful for our purposes here to tease apart two strands of the approach, and consider the arguments against them separately.

One, the principal one, is that the object of inquiry is behavior and that behavioral repertoires must be explained without reference to unobservable predicates. This last requirement, adopted from Logical Positivism, was observed more stringently by some people than by others, but it leads naturally to the restriction of the program of research in two directions. A psychological predicate like hunger, in the long run at least, has to be either "cashed out" in terms of behavior (or in the alteration of dispositions to respond to stimuli, presumably quantifiable in some way) or has to be replaceable by a set of, say, neurological and gastro-intestinal events which are correlated with hunger in a lawful way.

It appears, and I will take it for granted here, that both of these strategies have been in general unproductive, and that the arguments against them are compelling.⁴ Psychology, I will

⁴ See Dennett 1978, Fodor 1975, and Fodor and Block 1972, and the references they cite, for some discussion of this point, but suspicions like these have been around at least since Karl Lashley's famous address at the Hixon symposium in 1948. For some discussion of this, see Gardner 1985, especially chapter 2. Hilary Putnam, whose work will play an important role in what follows, was also influential in this development.

As a historical aside, we might note that there appears to be some disagreement as to the significance of Lashley's 1948 paper. Chomsky writes (in Language and Mind, p.3) that "his arguments and proposals, though sound and perceptive, had absolutely no effect on the development of the field and went unnoticed even at his own university (Harvard), then the leading center of psycholinguistic reserach. Ten years later, Lashley's

assume, must use predicates in its theory that are not, even in principle, either reduceable to natural (i.e. non-disjunctive) physiological predicates or mappable to behavioral predicates in a lawlike way. To use another terminology, psychology must make reference to "supervenient" properties. A supervenient property is one that must have a physical realization, but it needn't have the same physical realization in every thing that has the property (for a very clear discussion of this in the context of evolutionary theory, see Sober 1984). We might want to commit ourselves to the view that every psychological predicate has some physical realization or another, but the moral of this story is that focusing our attention on the physical realization is likely to be quite unilluminating, if what we're interested in is the nature of the human mind. Since there is no reason not to take the object of inquiry to be quite abstract, and in fact there are many reasons to take the object of inquiry to be quite abstract (because this will allow us to state the generalizations we're looking for), let's agree with Chomsky (and others) that the object of inquiry should be something like a cognitive structure, even though we have no idea what the physical realization of the cognitive structure might be. So this aspect of behaviorism, i.e. its conservative ontology, must be discarded, because it is counterproductively stringent.

The other strand of behaviorism, which by the way it shares with other theories of psychology (like Piaget's), is the interest in the laws of ontogenetic development of the mental

contribution began to be appreciated, but only after his insights had been independently achieved in other domains."

On the other hand, Gardner remarks (p.14) that "[F]rom comments by those attending the Hixon symposium, it seemed clear that Lashley's colleagues were deeply impressed by the originality and brilliance of [his] presentation... It is no exaggeration to suggest that entrenched modes of explanation were beginning to topple and that a whole new agenda was confronting the biological and behavioral communities." Perhaps this is merely a difference in emphasis, or in a different estimation of the significance of being ignored at Harvard in the 1950's.

that apply across tasks and across species. It really isn't a puzzle what Pavlov, Thorndike, and Skinner thought they could learn about humans by doing experiments with dogs, cats, and pigeons. The strategy was to try to eliminate the influence of the animal's biological propensities and its personal experience, that is, to isolate one of the various forces that impinge on a phenomenon, state what would happen if the force were acting alone (which may in practice never happen), and then describe actually occurring phenomena in terms of the interaction of these various forces. The strategy is as familiar as science itself. Gottlob Frege had a nice way of putting it: "When a problem appears to be unsolvable in its full generality, one should temporarily restrict it; perhaps it can then be conquered by a gradual advance" (1879:6).

So this is general process learning theory (a point of view that takes its departure from the work of Pavlov and Thorndike):⁵ It seeks to establish a set of laws that apply everywhere, regardless of the biological endowment, as we might say, of the organism and any particular organism's personal history. This is why the experiments had to be so, well, strange. If your goal is to discover general laws of learning, examining how a fish learns to swim (actually, fish probably don't learn how to swim in any

⁵ Here, and throughout this paper, I adopt the terminology used in Seligman and Hager's excellent introduction to their anthology (1972), to which I am indebted. However, I depart from these authors in two ways. First, for expository purposes, I state the equipotentiality premise in a different, but I think equivalent way. (They state it in terms of the technical vocabulary of classical and instrumental conditioning, which I want to avoid.) Second, it appears that I disagree with them on the basically historical question of how equipotentiality came to be regarded as a "fundamental premise" rather than as a research strategy (see below). To be honest, my reading of the relevant literature stems from an impulse to charity, rather than from a close analysis of the texts. Where I read the development as the result of a subtle, relatively late, but very important mistake, Seligman and Hager suggest that the mistake was made from the very beginning. Though I don't find their quotations to this effect convincing, perhaps they are right. At any rate, all are agreed that it was a mistake.

natural sense of the word "learn") is pointless. From this perspective, instances of one or no trial learning are not problems for the framework, they are simply irrelevant.

There is little question, I think, that many if not most psychologists working in the 50's and 60's did not see the issue this way, and this gives rise to what for me at least is a very puzzling episode in the history of the discipline. The question revolves around what came to be known as the "equipotentiality premise". We might put the premise this way: the genetic organization of an organism is irrelevant to learning. Of course, general process learning theorists were always trying to minimize the contribution of the organism's biology, in order to factor it out. This is a research strategy. In some circles, however, this research strategy was elevated to a fundamental truth, and the tail, so to speak, started wagging the dog. For if the equipotentiality premise is true, and general laws of learning are the only laws of learning that there are, one trial learning (like the acquisition of a taste aversion) becomes a problem, not an irrelevance.

I find it very difficult to think myself into this point of view. For the equipotentiality premise seems so very counter-intuitive and so easily falsified by the most mundane of observations. Be that as it may, it became necessary to falsify the premise in a way that would speak clearly to the psychologists of the day, for example, a carefully controlled experiment with rats (see Garcia and Koelling 1966).

These days, it's hard to find anyone who will admit that they ever believed that the premise was true in the first place.⁶

⁶ In a paper that makes very interesting reading, R.J. Herrnstein (1977) briefly surveyed the rise and fall of the equipotentiality premise. The source of the assumption is evidently difficult to locate. He doesn't name anyone who embraced the principle in print, writing:

While equipotentiality no doubt seems to be a common assumption among many Skinnerians, it is not to be

I prefer to read all of this as a sort of a technical correction in the marketplace of ideas. The bottom line is that if one wants to know about learning in general, the biology of an organism, as well as its personal history, are obviously relevant. Chomsky had a nice way of putting this at the end of his review of Verbal Behavior. He wrote (1959:578):

At any rate, just as the attempt to eliminate the contribution of the speaker leads to a "mentalistic" descriptive system that succeeds only in blurring important traditional distinctions, a refusal to study the contribution of the child to language learning permits only a superficial account of language acquisition, with a vast and unanalyzed contribution attributed to a step called generalization which in fact includes just about everything of interest in this process.

However, if one is trying to discover general laws of learning (if such exist) it makes sense to try to factor out the contribution of the organism's "preparation", be this biological or due to the organism's personal history.

2. So, where have we gotten to so far? We have rejected the view that the object of inquiry must be behavior, or indeed that the object of inquiry needs to be directly observable (whatever that might mean). So the thing we're trying to find out about can be quite abstract. We've agreed to call this a "cognitive structure" (whatever that might be). With everyone else, we've

found in Skinner's theoretical writings. What Skinner explicitly said about "representative" stimuli and responses involve "arbitrariness" not equipotentiality.

Skinner (1977) agrees. Arbitrariness of stimuli, of course, is just what one needs to have to explore general process learning theory within this framework. My hunch is that the smoking pistol lies in lecture notes discarded years ago. Evidently, some psychologists expressed points of view in lectures that were badly misleading. Thank heavens we've now overcome this impulse.

denied the equipotentiality premise, and observed that the organism itself makes a very important contribution to the acquisition of cognitive structures. The question that we're left with is this: do general laws of learning exist, and if so, what are they?

I'll now have a closer look at this question. Notice first, though, that it has a slightly different force now than it did for the learning theorists of yesteryear. For we're no longer asking whether there are general principles that regulate the acquisition of behavioral repertoires. We're asking whether there are general principles which regulate the acquisition of cognitive structures.

Chomsky has argued that the answer to the question is that there are no general laws of cognitive structure acquisition. This is the argument that I want to suggest fails. I'll then take up the question of what the implications of this are.

So far as I know, the issue first comes up in a response to a paper by Hilary Putnam (Putnam 1967), in which he suggested an affirmative answer to our question.⁷ We'll consider Chomsky's reply as it appears in Language and Mind (1972b:86):

Putnam does face this problem [the problem of how a

⁷ The suggestion comes in a context which, it seems to me, doesn't really give it the focus that it deserves. Putnam seems to imply that the existence of general multipurpose learning strategies would vitiate the argument for innate specification in the domain of language. Or at least he implies that we can make no judgment about innateness until we discover the general laws of learning.

I don't think any of this is convincing, but it does illustrate a point about science that was made by Quine (1951). Principles of a scientific framework are often regarded as not equal with respect to their falsifiability. If, like Putnam, you have a strong faith in general process learning theory, you'll tend to withhold judgements concerning principles more toward the "fringe", like innateness, on the suspicion that general process learning theory itself will explain what innateness is invoked to explain. On the other hand, if, like Chomsky, you have a deep belief in innateness, you'll tend to withhold judgments about general process learning theory, and (*mutatis mutandis*) for the same reason.

child acquires a grammar] and suggests that there might be "general multipurpose learning strategies" that account for this achievement. It is, of course, an empirical question whether the properties of the "language faculty" are specific to language or are merely a particular case of much more general mental faculties (or learning strategies)...

It's here, I think, that the confusion begins. We want to be able to ask whether there are general multipurpose strategies without confounding this question with the equipotentiality premise. So we want to grant that different species bring different "equipment" to the same tasks, and hence it is not at all surprising that the result of the application of this equipment is different. We want to grant this without prejudicing the question of whether or not there are regularities in how they apply the equipment they have.

Chomsky's way of setting up the issue doesn't encourage this. For a "faculty" is evidently intended to encompass both what the organism brings to a task and how the organism applies whatever it brings. Then a judgment concerning the specificity of a faculty won't tell us what we want to know. In fact, given our denial of the equipotentiality premise, we're probably prepared to grant that it's unlikely that any two different species will share the same faculty for any given task. So we could read Putnam as suggesting that some properties of the language faculty are general, and then again, maybe some are not. (For some further discussion of this point, see Flynn, to appear.) We will return to this point a little later. First, though, let's return to Chomsky's argument in Language and Mind (p.87):

If we discover through such investigation that the same "learning strategies" are sufficient to account for the development of competence in various domains, we will have reason to believe that Putnam's assumption is correct. If we discover that the postulated innate

structures differ from case to case, the only rational conclusion would be that a model of mind must involve separate "faculties", with unique or partially unique properties. I cannot see how anyone can resolutely insist on one or the other conclusion in the light of the evidence now available to us.

This needs to be sorted out a bit. I'll have to ignore the phrase "partially unique" since I'm not at all sure what Chomsky could mean by it in this context. The puzzler is the qualifier "innate" in the second sentence. I think it will help to adopt some terminology that was introduced by Chomsky later in Reflections on Language. During ontogenetic development, an organism proceeds from an initial to a final state in various domains. Innate structures of the initial state are not the issue here, since Putnam's proposal is meant to apply to the principles that regulate the organism's progression from initial state structures to final state structures. So far as I can see, "innate" here has to be read as an anticipation of an argument that Chomsky and Jerry Fodor would give years later (Chomsky and Fodor 1980 and Fodor 1980, but the argument was presented in 1975) to the effect that all mental structures have to be innate. Roughly, a final state can be distinct from an initial state and still be innate if the final state is a selection among the possibilities provided by the initial state. If we read this this way (and I don't see any other way to read it that will keep the passage relevant), what Chomsky's point amounts to is that the relation between initial states to final states (in various domains) is a function, in the mathematical sense. If we think of a "faculty" for a domain as the initial state for that domain and the operation that maps that initial state into a final state under maturation and experience, then if two final states are distinct, the associated faculties are distinct. That is to say, it never happens that one faculty maps the same maturational factors and experience into two different final states. This seems reasonable enough, but how it bears on general process

learning theory and Putnam's suggestion is still obscure.

To pursue the issue further, it will be easier to examine Chomsky's spelling out of this argument in Reflections on Language. The argument here is I think the same one that is hinted at in Language and Mind, except that now it is asserted to yield a conclusion: there are no general principles that regulate the acquisition of cognitive domains.

It is set up with the aid of the neutral scientist S, a figure who by now is quite familiar to linguists. This person is unencumbered by pretheoretical conceptual baggage, uninfluenced by tradition, received opinion, promotion possibilities, political ideologies and funding sources. S is interested only in truth.

Chomsky writes (1975:14):

Consider first how a neutral scientist - that imaginary ideal - might proceed to investigate the question [i.e. the question of whether or not there is a general theory of learning]. The natural first step would be to select one organism O and a reasonably well-defined cognitive domain D, and attempt to construct a theory that we might call "the learning theory for the organism O in the domain D." This theory - call it $LT(O,D)$ - can be regarded as a system of principles, a mechanism, a function, which has a certain "input" and a certain "output" (its domain and range respectively). The "input" to the system $LT(O,D)$ will be an analysis of the data in D by O; the "output" (which is, of course, internally represented, not overt and exhibited) will be a cognitive structure of some sort.

$LT(H,L)$, for example, is the learning theory for humans in the domain of language, which is "the system of principles by which humans arrive at a knowledge of language, given linguistic experience, that is, given a preliminary analysis that they develop for the data of language." $LT(R,M)$, on the other hand, is the learning theory of rats in the domain of maze running.

So $LT(O,D)$ is what a moment ago I was calling O's D faculty, that is, it is a function which maps O's pretheoretical analysis of the data in D to the cognitive structure which O achieves in D. It evidently consists of (perhaps among other things) what O innately "knows"⁸ about D together with some instructions for what to do with this knowledge and the relevant experience in D, i.e. how to go about building the cognitive structure for D. Now we can re-ask the question about general process learning theory (this is Chomsky's question (1) on p.17):

Is it the case that however we select O and D, we find the same $LT(O,D)$?

Chomsky states that the answer to this question must be "a firm No." He continues,

Even the crudest considerations suffice to show that there is no hope of reaching a positive answer to this question... Even if some vague approximation to [this question] had a positive answer, we would expect humans to be as much superior to rats in maze-learning ability as they are in language-learning ability. But this is so grossly false that the issue cannot be seriously entertained...

To take Chomsky's examples, suppose we consider four cognitive structures achieved by organisms:

Language-learning in humans:	CS_{hl}
Language-learning in rats:	CS_{rl}
Maze-running in humans:	CS_{hm}
Maze-running in rats:	CS_{rm}

While CS_{hm} and CS_{rm} are roughly comparable, CS_{hl} and CS_{rl} are

⁸ I will henceforward drop the scare quotes around this and other words which in other contexts would beg important questions. To digress here to discuss these issues or to invent a specialized vocabulary would only be distracting and add to the tedium.

very different. From this, it is alleged to follow that the corresponding LT's must be different. General process learning theory can be easily seen not to exist, and from this perspective it is hard to see why anyone would have ever thought it existed.

However, I believe this argument is mistaken, in that it doesn't show anything at all about (general process) Learning Theory. It's easier to see this if we make the notation a bit more perspicuous and make explicit an important part of the argument.

Recall that LT's like $LT(H,L)$ and $LT(R,M)$ are functions from data to cognitive structures. Don't be misled by the notation here. It is not intended that LT is a function from pairs like (H,L) and (R,M) to cognitive structures, but rather that an indication like (H,L) tells us which LT we are considering. So we might instead write LT_{hl} and LT_{rm} .

Now here is an important point. Recall that LT_{od} is a function from O's "pretheoretical characterization of the data in D". I take it that everyone agrees that this will vary from organism to organism. This is just the denial of the equipotentiality premise, and is why general process learning theorists never thought to study, for example, the acquisition by birds of the ability to fly. We've seen that this is independent of the question of general process learning theory. So, in order to keep the strands of the issue untangled, let's make an assumption of instantaneous acquisition (on this see Chomsky 1975:119ff) and let whatever unique, special and biologically determined knowledge the organism brings to the task result in the differences in which different organisms pretheoretically analyze the data. Let's call the pretheoretical analysis by O of the data in D \mathcal{D}_{od} .

So here is the situation. We have:

$$LT_{hl}(\mathcal{D}_{hl}) = CS_{hl}$$

$$LT_{rl}(\mathcal{D}_{rl}) = CS_{rl}$$

$$LT_{hm}(\mathcal{D}_{hm}) = CS_{hm}$$

$$LT_{rm}(\mathcal{D}_{rm}) = CS_{rm}$$

Looking at the issue this way I think reveals the following observation: if we agree that all the CS's are different (as we do) and we agree that all the \mathcal{L} 's are different (as I think we must) then nothing whatsoever follows about the relative character of the LT's. In particular, there is no contradiction in supposing the LT's to be identical, which can be seen by substituting f for all of them above and observing that it is obviously not intended that f be a constant function (i.e. a function which returns the same value at all arguments). On the other hand, the LT's could all be distinct as well. I conclude from all of this that the argument in Reflections on Language doesn't bear at all on the question of general process learning theory, and that the agnosticism on the issue expressed by Chomsky in Language and Mind is quite appropriate.

A word of caution. When pondering matters like this, it is easy to set up the situation so that general process learning theory and the equipotentiality premise rise and fall together. For example, in the present case, we could have assumed that all the \mathcal{L} 's were identical, and attributed the biological contribution of the organism to the LT's. Here, making the (quite reasonable) assumption that the LT's are functions leads immediately to the conclusion that they are distinct. But this isn't very helpful if we're already prepared to deny equipotentiality, as, of course, we are. If we want to know about some principle, it's important to set up our investigation so that we keep that principle independent of other principles we've already made a decision about.

3. So, as advertized, we have gotten nowhere on the general process learning theory question, but it may be of some comfort to be aware of this. While we're here, though, it might be worth lingering a bit to ask ourselves what the thing would be like if it were to exist. This question can't be sensibly asked in isolation since, as we've already seen, the force of the question varies depending on what other properties we think the mind has.

So let's make some assumptions. Suppose organisms are innately endowed with highly articulated principles of mental organization, some of which are specific to a particular domain. We'll call this the initial state of the organism. (We'll leave the concept of a domain fuzzy, as perhaps it should be. On this matter, see Flynn, to appear.) We'll further assume that under the push and pull of the environment the organism achieves (if all goes well) various coherent and identifiable final states, and that these final states are cognitive structures whose character is determined by the setting of parameters offered to the organism by its biology. We want to know (among other things) the principles that regulate this achievement, keeping in mind that it's quite possible that these principles vary from domain to domain.

One thing to note is that it isn't necessary that all or even any of these principles are represented in the mind, in the sense that the principles are subject to mental computations. (See Fodor 1975 on the relation between representation and computation.) In fact, I think it's possible to read Carroll 1895 and Godel 1931 as implying that they can't all be represented. Be that as it may, it seems clear, given the other assumptions we've made, that the principles that we're after will be akin to rules of inference, in the sense that what they do is license the organism to move from one state to another.

To appreciate the logic of the situation, let's boldly (if crudely) speculate for a moment. Suppose we find some domain in which we discover we have some reason to believe that one of the principles that allow an organism to advance from one state to the next is in fact one of the rules of inference we all came to know and love in logic class. Suppose it's modus tollendo ponens, the rule which licenses an inference like this:

A or B
not A
therefore, B.

If this were to happen, I think we would feel fairly comfortable

in suspecting that at least one principle which allows an organism to move from one state to another is a fairly general one, i.e. that at least one principle plays a role in more than one domain. Which is just to say that we would have grounds for believing that general process learning theory exists. In other words, we would be adopting a position that while some principles that comprise faculties are domain (and species) specific, others are general, i.e. they play a role in more than one faculty.

We might begin to suspect that at least some other species achieve what cognitive structures they achieve in part by means of a similar process. I've heard it confidently asserted that some household pets are extremely good at modus ponens. In fact, classical conditioning seems to be predicated on this. I'm not sure what to make of this now, since I don't know of any studies of other species that have been conducted under the assumptions that we are assuming. But I do think that, from this point of view, it is possible, if not plausible, to entertain the following scenario:

Humans and, say, pigeons, do share aspects of their cognitive organization (much like they share some basic aspects of their respiratory and circulatory systems), namely, the method by which their cognitive structures are acquired. The reason the two species achieve different cognitive structures is that they start out in different places. With Chomsky, we correctly balk at the extension to humans of a theory of psychology worked out by studying other animals. But we should be less hesitant about extending to other animals a theory of psychology worked out by studying humans.

With this in mind, consider the following passage from Richard Kayne's very interesting paper "Extensions of Binding and Case-Marking" (chapter 1 in Kayne 1983:12).

Let us recast the reasoning into terms of the learner constructing a grammar of French. He learns that French has a Rightward NP Movement rule and general principles give him the derived structure of that rule,

at least to the point of ensuring that the moved NP is not a proper binder for its trace. The NIC tells him that such a configuration is illicit. Yet he knows that (61) is well formed.⁹ Consequently, he must alter the surface structure of (61) to bring it into accord with the NIC. Thus he postulates a rule that puts the already displaced subject NP into a position that does properly bind its original trace.

Of course, the details, indeed even the specific content of this passage, are not so important for our purpose here. We're interested in the form of this story, and the form of the story is that of a (fairly complicated) logic problem. I think that Kayne is presupposing that among the problems faced by the little language learner are problems which require deductive logic to solve, and I suspect he's right. Returning to our modus tollendo ponens (MTP) case above (because it's simpler), if universal grammar says that languages have either property A or property B (or both) and the child discovers that it doesn't have property A (perhaps a supplementary principle says that if a language has A at all, it shows up in very simple sentences), then, if we assume the acquisition device has MTP as one of its components, the inference is made that the language must have property B, even though the child has perhaps not yet seen any instances of B. (For another example that I think can be slightly modified to exemplify this perspective, see Williams 1981.) I don't think we need worry too much about the exact form of the inference rules employed until we have reason to believe that they are mentioned instead of just being used in mental computations.

There is a subtlety here which we ought to take note of. (I thank Paul Saka for bringing this to my attention.) There is a sense in which logic is necessarily involved in all scientific

⁹ [(61) is
le jour ou sont arrives beaucoup de garcons
the day when have arrived many boys
(i.e. when many boys arrived)]

theories, if we require of scientific theories that they have deductive structure (as we do). Logic does form a component in how we go about reasoning about the world, including our reasoning about how organisms construct the cognitive structures that they do. It doesn't follow from this, however, that principles of logic form a substantial component of the acquisition process. In our example, it is by no means necessary that a child infers the existence of the property B without having seen any instances of it. What's crucial is the assumption that the child antecedently knows that the language will have A or B or both, since without this assumption, it's hard to see what role MTP could play for the child, though it does, of course, play a role in our theory of what has happened. It's only the (empirical) assumption of a rich and highly articulated biological endowment that makes it at all plausible that the principles of logic play a crucial role in cognitive structure acquisition.

Let me finish with what I think is the main point. Principles in a theory interact in a number of ways. We've seen several of these in this paper. Two principles can be independent in the sense that the falsification of one doesn't imply either the truth or the falsity of the other. This is the case with the equipotentiality premise and general process learning theory. Nevertheless, they can be so closely associated in the minds of the theorists, that their popularity tends to rise and fall together. This is a sociological matter, though that doesn't make it uninteresting, of course. I think this happened in the minds of many as the light of behaviorism began to flicker in the 50's and 60's.

Principles can also be related in that a change in one requires, or at least invites, a change in another. In our case, when you change the object of inquiry away from the acquisition of behavioral repertoires and towards the acquisition (or growth, if you like) of cognitive structures, it seems sensible to modify the question of the existence of general process learning theory so that the spirit of it can be applicable to investigations

within the new theoretical ambience.

Another way principles interact, and as an interested observer of the scene I suppose I find this the most intriguing, is that sometimes, when one framework gives way to another, some of the principles of the first gain, rather than lose, plausibility. As Jerry Fodor (1980) has forcefully, and I think persuasively, argued, it's very difficult to see how any general theory of induction could explain the ability of a child to develop rich mental structures.¹⁰ It's hard to be optimistic about general process learning theory when looking at a framework like Piaget's. However, we know a lot more about deduction. When we re-ask the question in a framework which sees the organism deducing aspects of its mental organization from premises supplied by its biology and experience, stock in general process learning theory would seem to rise a bit. But unlike on Wall Street, there's no insider trading in the psychology market that I know of, so caveat emptor.

¹⁰ There is here potential for some confusion concerning an issue that strikes me as terminological. Suppose someone says to me, "Look, what you're talking about can't be learning, because learning is by definition inductive." Well, maybe so. It's in this sense that Fodor can claim with some justification that "there is no learning theory [and] in certain senses there couldn't be" (1980:143). But the view that I'm entertaining here is that there is a component of the learning process which is deductive, and this component is what, or part of what, instances of learning share across species and across tasks. None of the conclusions that the learner draws, however, will be demonstrable, because in general the truth of the premises will not be demonstrable.

In our example with MTP, we will in the end have to say something about how the language learner comes to the decision that the language doesn't have the property A. Evidently, some kind of nondeductive inference is involved here. Now, we could describe what has happened this way: the child has learned that the language fails to have A, but the child has not learned that the language has B, since the second follows deductively from Universal Grammar and the first inference. This seems to me to be an unnatural way of describing the event. I would rather say that the child has learned that the language has the property B, and that one step in the process involved deduction. It comes down to how we want to use the word "learn".

I said my conclusion would be modest, and I think it is. I've urged that there's no reason to think that general process learning theory is only an illusion, and that there's some reason to think that it is something to investigate when we turn to wondering about the structure of the human mind. Behaviorism may have been a disaster, but I don't think it was an unmitigated one.

References

- Block, Ned and Jerry A. Fodor. 1972. "What Psychological States Are Not", Philosophical Review 81.2:159-181. Reprinted in Fodor 1981.
- Bloomfield, Leonard. 1926. "A Set of Postulates for the Science of Language", Language 2:153-164. Reprinted in Readings in Linguistics I, edited by Martin Joos. University of Chicago Pr, 1957.
- Borges, Jorge Luis. 1964. Labyrinths. New Directions Publishing, New York.
- Carroll, Lewis. 1895. "What the Tortoise said to Achilles", Mind 4:278-280. This paper has been frequently reprinted. One place is in Howard DeLong, A Profile of Mathematical Logic. Addison-Wesley Publishing, Reading MA, 1970.
- Chomsky, Noam. 1959. "A Review of B.F. Skinner's Verbal

- Behavior", Language 35.1:26-58. Reprinted in The Structure of Language, edited by J.A. Fodor and J.J. Katz. Prentice-Hall, Englewood Cliffs NJ, 1964.
- Chomsky, Noam. 1972a. "Psychology and Ideology", Cognition 1.1: 11-46.
- Chomsky, Noam. 1972b. Language and Mind. Harcourt Brace Jovanovich, New York.
- Chomsky, Noam. 1975. Reflections on Language. Pantheon Books, New York.
- Chomsky, Noam and Jerry Fodor. 1980. "Statement of the Paradox", in Piattelli-Palmarini.
- Dennett, Daniel. 1978. Brainstorms. Bradford/MIT Pr.
- Flynn, Michael. To appear. "Categorial Grammar and the Domain Specificity of Universal Grammar", in Modularity in Knowledge Representation and Natural Language Understanding, edited by Jay L. Garfield. Bradford/MIT Pr.
- Fodor, Jerry A. 1975. The Language of Thought. Thomas Y. Crowell Company, New York.
- Fodor, Jerry A. 1980. "Fixation of Belief and Concept Acquisition", in Piattelli-Palmarini.
- Fodor, Jerry A. 1981. Representations. Bradford/MIT Pr.
- Frege, Gottlob. 1879. Begriffsschrift, in van Heijenoort 1970.
- Garcia, John and Robert A. Koelling. 1966. "Relation of Cue to Consequence in Avoidance Learning", Psychonomic Science 4: 123-124. Reprinted in Seligman and Hager 1972.
- Gardner, Howard. 1985. The Mind's New Science. Basic Books, New York.
- Godel, Kurt. 1931. "On Formally Undecidable Propositions of Principia Mathematica and Related Systems I", in van Heijenoort 1970.
- Herrnstein, R.J. 1977. "The Evolution of Behaviorism", American Psychologist 32.8:593-603.
- Kayne, Richard S. 1983. Connectedness and Binary Branching. Foris, Dordrecht.
- Lyons, John. 1978. Noam Chomsky. Penguin, New York.
- Piattelli-Palmarini, Massimo (ed). 1980. Language and Learning: The Debate between Jean Piaget and Noam Chomsky. Harvard UP.
- Putnam, Hilary. 1967. "The Innateness Hypothesis and Explanatory Models in Linguistics", Synthese 17.1.
- Quine, W.V.O. 1951. "Two Dogmas of Empiricism", Philosophical Review 60:20-43. Reprinted in W.V.O. Quine, From a Logical Point of View. Harvard UP, 1964.
- Seligman, Martin E.P. and Joanne Hager (eds). 1972. Biological Boundaries of Behavior. Appleton-Century-Crofts, New York.
- Skinner, B.F. 1977. "Herrnstein and the Evolution of Behaviorism", American Psychologist 32.12:1006-1012.
- Sober, Elliot. 1984. The Nature of Selection. Bradford/MIT Pr.
- van Heijenoort, Jean (ed). 1970. Frege and Godel. Harvard UP.
- Williams, Edwin. 1981. "X-bar and the Analysis of Categories", in Language Acquisition and Linguistic Theory, edited by Susan Tavakolian. MIT Pr.