

## EMPIRICISM, INNATENESS, AND LINGUISTIC UNIVERSALS

(Received 21 May, 1977)

For the last decade and more Noam Chomsky has been elaborating a skein of doctrines about language learning, linguistic universals, Empiricism and innate cognitive mechanisms.<sup>1</sup> My aim in this paper is to pull apart some of the claims that Chomsky often defends collectively. In particular, I want to dissect out some contentions about the existence of linguistic universals. I shall argue that these claims, while they may be true, are logically independent from a cluster of claims Chomsky makes about Empiricism, language learning and innate cognitive mechanisms. The latter claims are, on my view, significantly more plausible than the former. Failure to note the logical space between the two has had lamentable consequences both for Chomsky and for his critics. Chomsky has been rather too ready to accept the existence of linguistic universals, and in arguing that one or another specific feature of a grammar is universal, he has consistently invoked an argument form of dubious merit. The critics, for their part, have sometimes argued against Chomsky's attack on Empiricist learning theories by criticizing Chomsky's views on language universals. If, as I shall maintain, the two positions are quite independent, then the critics' skepticism about linguistic universals is largely irrelevant to the Chomskian critique of Empiricist theories of learning.

To begin, let me sketch those parts of Chomsky's view that I take to be most plausible. The matter at hand is language acquisition and the cognitive mechanisms that underlie it. What is it that happens when a child learns her first language? According to Chomsky, she comes to know, albeit tacitly, the rules of the grammar of her language. A number of philosophers, myself among them, have badgered Chomsky about this way of viewing language learning.<sup>2</sup> In what follows, however, I propose to ignore those objections and adopt Chomsky's view that in acquiring her first language a child internally represents a grammar of her language, or, what is essentially the same view, that she comes to tacitly know the rules of her grammar.

*Philosophical Studies* 33 (1978) 273–286. All Rights Reserved  
Copyright © 1978 by D. Reidel Publishing Company, Dordrecht, Holland

Granting this view of language learning, what can we say about the psychological mechanisms that are responsible for the child's feat? It could hardly be denied that for learning to be possible the child must come to the task already equipped with some innate learning mechanism.<sup>3</sup> But agreement on this point leaves plenty of room for disagreement. Perhaps the most exciting question in dispute is this: Could 'Empiricist' mechanisms — the sorts of mechanisms characteristic of Empiricist theories of mind and learning — possibly be sufficient to explain the process of language acquisition; or must we instead posit some very different sort of mechanism, one which is more Rationalist in cast? Chomsky, of course, holds that an Empiricist learning theory cannot account for language learning.<sup>4</sup>

There is, in Chomsky's writings, a variety of arguments aimed at establishing the inadequacy of Empiricist learning theories. By my lights, the most persuasive of these are the various versions of what might be called *the rational scientist argument*. In outline, the argument goes like this: Suppose an empirical investigator were to be set the task of duplicating the child's achievement. The investigator is given all the data available to the child during language acquisition, and only those data. The investigator is assumed to have no explicit knowledge about universal features of human language, if such there be. Would such an investigator be able to discover the grammar of the language to which the child is exposed? The issue, of course, is not whether the investigator might be able to fortuitously think up the correct grammar, but rather whether he would be able to be tolerably sure that he had the right one, given the evidence available. The answer, Chomsky contends, is *no*. For there are indefinitely many grammars all compatible with the evidence our imagined investigator has available, yet significantly different from each other. Moreover, there is no reason to believe that the correct grammar will be simpler than its competitors on any intuitive, non-question begging account of simplicity. So the investigator would not be able to infer from the evidence the child has available to the grammar the child in fact acquires. Now to tie this conclusion to the attack on Empiricism we need only note that our rational investigator presumably has available to him all the inferential apparatus compatible with Empiricist theories of learning, and indeed he has a bit more.<sup>5</sup> So if he cannot infer from data to grammar, no Empiricist acquisition mechanism can duplicate the child's trick of coming up with the correct grammar on the basis of the data available.

The rational scientist argument, it need hardly be said, is a topic of

considerable controversy. My own view is that, though it plainly needs much elaboration, the argument as sketched is the core of a very persuasive case against the adequacy of Empiricist language acquisition devices. However, it is no part of my current project to defend the rational scientist argument. I want to quarrel with Chomsky, not to praise him. So let us agree, even if only for argument's sake, that Chomsky has shown the correct language acquisition theory will be a non-Empiricist theory. The quarrel I have with Chomsky is that he thinks he has shown quite a bit more. Chomsky has long advocated the view that the correct language acquisition theory will be a non-Empiricist theory of a very special kind. So special, indeed, that if he is right, then the existence of linguistic universals follows straightforwardly. The burden of my argument is that there is neither evidence nor argument in favor of the view that the correct acquisition theory must be of this very special sort. Such argument as we are given would at most establish that the right acquisition theory is a non-Empiricist one. My hunch is that Chomsky has simply failed to note the diversity possible in non-Empiricist acquisition theories, and thus has failed to see the gap between the defended view that a correct acquisition theory will be non-Empiricist, and the undefended view that it will be the kind of theory which entails the existence of linguistic universals. The first step in making my case is to underscore the potential for diversity in non-Empiricist theory of language acquisition. To this end I will sketch four categories of acquisition theory. The theories in each category plainly attribute more to the mind, by way of innate endowment, than would any theory that might comfortably be called Empiricist.

*Type I.* Theories in this group explain language acquisition by maintaining that all the details of the grammar the child will come to know, and only the details of that grammar, are already internally represented in the infant at birth. Time is needed for the innately programmed grammar to be properly utilized, thus the child does not speak at birth. But the process of 'learning to speak' is simply a maturational process, like puberty or the growth of teeth. Theories which attempt to account for language learning in this way have the modest virtue of being unquestionably false. For it is plain that the child's external environment during the time she passes from non-speaker to speaker is not as irrelevant to determining the end product of acquisition as this theory suggests. Regardless of parentage, children raised in an English speaking environment internalize English grammar, while children raised in a Chinese environment internalize Chinese grammar.

*Type II.* A second group of acquisition theories are those which fit what might be called the Innate Library Model. On theories of this sort, the child is viewed as innately having internal representations of every grammar it is possible for a human to learn. The representations are complete in every detail; each contains all the grammatical information an adult speaker of the language would tacitly know. Such theories will also have to postulate an innate search mechanism of considerable complexity. What is needed is some efficient way of selecting the correct grammar from the mass of incorrect ones, on the basis of the relatively limited cues provided by the input to the acquisition mechanism. The possible heuristic strategies for seeking the correct grammar are limitless in number; some might invoke a system of ranking of grammars (and of categories of grammars), others would not.

There is a distinction to be drawn between two rather different sorts of Innate Library Model theories which will play a central role in our later argument. The library metaphor suggests a useful first pass at explaining the distinction. Imagine two different sort of libraries: The first sort of library, like most libraries, has a wide variety of books, some of which may be strikingly similar to each other (e.g., consecutive editions of the same title) and some of which have none but the most trivial properties in common with the other books in the library. Quite a different sort of library is one whose collection is limited to closely related volumes. An extreme case might be a library that collected only books with identical first chapters. Another example might be a library that specialized in stories of a certain quite definite structure with a specific sequence of events, though the characters might have different names and the events might take place in different places. This distinction between restricted and unrestricted libraries applies fairly directly to acquisition theories that fit the Innate Library Model. Restricted Library Model theories attribute to the child innate knowledge of a class of grammars that have non-trivial properties in common. Unrestricted Innate Library Model theories attribute to the child innate knowledge of grammars whose only common properties are utterly trivial ones. The distinction is hardly a precise one, since I have not said what I take to be trivial properties. For the most part, I am content to rely on our shared intuitions about which common properties are trivial and which are not. But perhaps a few illustrations may serve to sharpen the reader's understanding of what I mean by a 'trivial' common property. Presumably the innately known grammars are finite in number. Suppose them to be en-

numerated in some arbitrary way. Then each grammar in the library has the property of being identical with the first, or being identical with the second, of ..., or being identical with the last. This is a paradigm of the sort of common property I am calling 'trivial'. Common properties of this sort can be found in any finite set. Another trivial common property might be containing no more than  $n$  rules, for some suitably large  $n$ . For future references, let us call *unrestricted* Innate Library Theories *type II-U theories*, and *restricted* Innate Library Theories *type II-R theories*.

I know of no argument against type II theories that is quite as knock-down as the argument against type I theories. But still it seems wildly implausible that a type II acquisition theory could be correct. There are thousands of distinct languages known, and no doubt there are thousands more humanly possible languages that are not now and never have been spoken. Since each language has a distinct grammar, the innate storage capacity that would have to be postulated by a type II theory is simply staggering.<sup>6</sup> So it looks like we will have to look elsewhere for plausible theories of language acquisition.

*Type III.* The third group of acquisition theories attributes to the pre-linguistic child not an innate library of grammars, but rather an innate *production system*, or set of rules, which generates grammars. Among the grammars generated by the rules will be all and only those which can be acquired by a human being in the natural way. Perhaps a word of explanation is in order on the notion of a production system that generates grammars. In general terms, a grammar-generating system is analogous to a grammar. A grammar is system of rules which generates a large number of symbol sequences; if it is a grammar of English, then each of the symbol sequences generated will be (a phonetic representation of) an English sentence, and (a phonetic representation of) each English sentence will be generated by the grammar. Now a grammar itself is simply sequence of symbols, though of course it is a much longer and more complex sequence of symbols than most commonplace sentences. So just as there may be a system of rules which would generate each English sentence, so too there may be a system of rules which would generate each humanly learnable grammar. According to type III theories, normal children are born with a grammar-generating production system which they exploit in learning the grammar they do.

As with type II theories, type III theories will have to attribute more to

the child than merely the innate grammar-generating production system. The child will also have to have some very sophisticated heuristic mechanism which enables her to exploit sensory input at various points in using the production system. Put in another way, the job of the heuristic mechanism will be to generate the correct grammar from the production system, using the sensory input available to the child as guidance.<sup>7</sup>

The distinction between unrestricted and restricted Innate Library theories (i.e. between type II-U and type II-R theories) finds an exact analogue in Innate Production System theories. There are some production systems whose output grammars will all have significant and non-trivial properties in common, while other production systems will generate a heterogeneous class of grammars whose only common properties are trivial ones. It might be thought the very fact that a class of grammars is generated by a given production system is sufficient to guarantee that each of the generated grammars will share non-trivial properties. But this is not the case. Perhaps the simplest way to make the point is this: Suppose we have 17 arbitrarily different production systems, each generating a different class of grammars. Further, suppose that the first rule of the first production system is of the form  $G_1 \rightarrow \text{---}$ , the first rule of the second production system is of the form  $G_2 \rightarrow \text{---}$ , ... and the first rule of the 17th production system is of the form  $G_{17} \rightarrow \text{---}$ . Then we can construct a new production system whose output class is the union of the output classes of the 17 hypothesized systems. We need simply take the first rule to be:<sup>8</sup>

$$G \rightarrow \left\{ \begin{array}{c} G_1 \\ G_2 \\ \vdots \\ G_{17} \end{array} \right\}$$

Now plainly there is no reason to suppose that an arbitrary grammar generated by the 17th production system has any but trivial properties in common with an arbitrary generated by the first production system. But both of these grammars are generated by our larger production system. So we have at least one straightforward example of a grammar-generating production system whose output grammars share only trivial properties. More generally, there is no more reason to suppose that the output grammars of a grammar-generating production system will have non-trivial properties in common

than there is to suppose that the output sentences of a grammar will have non-trivial properties in common. Let us call those Innate Production System theories whose output grammars share only trivial properties *unrestricted Innate Production System theories* (or *type III-U theories*), and let us call those whose output grammars share non-trivial properties *restricted Innate Productions System* (or *type III-R theories*).

*Type IV.* The final category of acquisition theory I will mention is a variation on type III theories. In type III theories we supposed that the child was endowed with an innate production system which generated complete, fully detailed grammars. Type IV theories, by contrast, attribute to the child an innate production system which generates only partially complete grammars. To fill out the details of the grammar, the child invokes one or another essentially empiricist mechanism. Thus, for example, one type IV theory might generate grammars which were fully detailed save for not specifying the phonetic shape of items in the surface vocabulary. Discovering the phonetic shape of surface vocabulary would be relegated to some sort of empiricist mechanism which could 'copy' the correct phonetic information from the speech the child heard. Of course, like type III theories, type IV theories will also have to attribute to the child a sophisticated heuristic mechanism whose function is to use environmental input to guide the production system in generating the correct output. The boundary between type III and type IV theories, like that between type II and type III, is a fuzzy one, and it is no trick to imagine acquisition theories which might be counted as either type III or type IV. Type IV theories, like type III, can be further divided into two sub-types, with the distinction turning on the matter of common properties in the class of output grammars. If all the (partially complete) grammars generated by a type IV theory have non-trivial properties in common, we will call the theory a *type IV-R theory*; if the output grammars share only trivial properties, we will call the theory a *type IV-U theory*.

This completes my brief taxonomy of non-Empiricist theories. I do not offer it as a complete catalogue, for no doubt there are quite different sorts of non-Empiricist language acquisition theories that might be constructed. Rather, my purpose was to point out just how diverse non-Empiricist theories can be, and to help us locate the sort of theory Chomsky advocates. Just what kind of language acquisition theory does Chomsky think is correct? The answer is not entirely clear. In *Aspects of the Theory of Syntax* Chomsky insists that "a child who is capable of learning a language must have ... (iii)



some initial delimitation of a class of possible hypotheses about language structure ... [and] (v) a method for selecting one of the (presumably, infinitely many) hypotheses that are allowed by (iii) and are compatible with the given primary linguistic data." (p. 30) These remarks, along with most everything else that Chomsky has written about theories of language acquisition are compatible with the proposal that the correct theory will be of type III or type IV. Chomsky's remarks do not commit him to a type III or type IV theory since he is generally non-committal on just how the "initial delimitation of [an infinite] class of possible hypotheses about language structure" is to be accomplished. Thus there may be some sort of acquisition theory which does not delimit the class of possible output grammars by using a production system, but does it instead in some other way. For the sake of explicitness, however, I shall in what follows assume that Chomsky would expect to find the correct acquisition theory among the type III or type IV theories. Little generality is lost here since the point I want to make can almost certainly be reconstructed for most any plausible alternative account of how the output class of the acquisition theory is specified.

Theories of types III and IV are surely the most plausible among the acquisition theories we have surveyed. Thus I have no quarrel with Chomsky's view that the correct theory is to be found here. However, Chomsky also insists that the correct theory will be a *restricted* or *type R* theory, and it is here that my dissent is focused. Chomsky insists that an explanatorily adequate linguistic theory

incorporates an account of linguistic universals, and attributes tacit knowledge of these universals to the child. It proposes, then, that the child approaches the data with the presumption that they are drawn from a language of a certain antecedently well-defined type, his problem being to determine which of the (humanly) possible languages is that of the community in which he is placed. Language learning would be impossible unless this were the case. (*Aspects*, p. 27).

Moreover, it is clear from what Chomsky says in the following text and in many other places that the universals he is speaking of are non-trivial ones. In particular, he is plainly claiming *more* than that all humanly possible grammars share the common property of being generated by *some* particular grammar-generating production system. So Chomsky is committed to the claim that the output class of the acquisition mechanism, i.e. the class of 'humanly possible grammars', will exhibit non-trivial common properties. Now the problem with this claim is that *it is entirely undefended*. Chomsky



does not offer any serious empirical evidence for the existence of non-trivial linguistic universals.<sup>9</sup> The arguments he does offer are versions of the rational scientist argument. But this argument at best establishes that an acquisition theory must be non-Empiricist. It is irrelevant to the question of linguistic universals unless it is assumed that any non-Empiricist theory must postulate linguistic universals. And as we have seen, that is simply false. Type III-U and type IV-U theories are clearly non-Empiricist, though neither category of theory insists that all humanly possible grammars exhibit non-trivial universal features.

Perhaps the most striking symptom of Chomsky's failure to note the logical space separating non-Empiricist learning theories from the doctrine of linguistic universals is his fondness for a curious form of argument. When arguing that one or another feature of a grammar is in fact universal to human grammars, Chomsky often will rest his case on the observation that the feature in question could not have been learned by the child. To show that it could not be learned, he points out that there are numerous alternative logically possible grammars which lack the putatively universal feature but which are compatible with all the evidence available to the child. Since the child acquires a grammar which exhibits the feature, Chomsky argues, the feature must be universal.<sup>10</sup> Now it takes relatively little to see that the conclusion in Chomsky's argument is simply a non-sequator. From the fact that the evidence available to the child is compatible with grammars not exhibiting the putatively universal feature, we might plausibly conclude that the acquisition mechanism itself is heavily implicated in the explanation of the fact that the child acquires a grammar exhibiting the feature. But it is quite another matter to claim that the feature is universal among the possible outputs of the acquisition mechanism. For Chomsky's argument to be at all plausible in establishing this latter conclusion, we should have to assume not only that there are linguistic universals, but also that most (or all) of the features of grammars which are unlearned (i.e. which the child (or the rational scientist) cannot inductively discover from the evidence available) are universal features.

It might be thought that the universality of unlearned features could be established as follows:

Since the evidence available to the child is compatible with alternative (logically possible) grammars which do not exhibit the feature in question, nothing in the data can tell the acquisition mechanism to choose a grammar with the feature rather than a grammar

without it. The fact that the acquisition mechanism does select a grammar with the feature can only be explained if we assume that it has no choice, because all of its possible outputs have the feature.

However, this argument turns on a confusion. In granting that the child's evidence is *compatible* with alternative logically possible grammars not exhibiting the feature, we are surely *not* granting that "nothing in the data can tell acquisition mechanism to choose a grammar with the feature rather than a grammar without it". We are granting only that nothing in the data lends *inductive support* to the choice of one grammar over the other. Put more vividly, we are granting that the rational scientist would be unable to select between a pair of grammars, one exhibiting the feature and one not, solely on the basis of that data available to the child. But the acquisition device need not be similarly indecisive. Its decisions are not constrained by the canons of inductive reasoning. They may turn on some quite *arbitrary* aspect of the available data, an aspect which the scientist would rightly view as evidentially irrelevant to the choice confronting him. Thus, to conjure a slightly outlandish example, it might be that the acquisition device outputs 17 quite distinct categories of grammars and that the members of only one of these categories, say the 13th, exhibit a certain abstract structural feature. The heuristic mechanism, in turn, might be innately programmed to restrict its search to category 13 if and only if, say, the word commonly used for a certain shade of red in the child's linguistic community is polysyllabic. Plainly the argument we have been reviewing lends no support to the claim that unlearned features of a grammar need be universal.

Thus far I have argued that an attack on Empiricist acquisition theories is no defense of the doctrine of linguistic universals. Before concluding let us ponder just what sort of defense that latter doctrine might receive. The most obvious sort of evidence to be sought for the claim that all humanly possible languages share common features would come from an exhaustive survey of extant languages. If we find non-trivial common features in the grammars of large numbers of diverse languages (and, of course, if we find no language whose grammar lacks the feature in question) then we have some serious evidence for the claim that there are linguistic universals. But in syntax and semantics, at least, it seems dubious that any such evidence will be in the offing soon, since there is precious little agreement about what the correct generative grammar of even the most carefully studied languages will look like.

Even if there were considerable evidence that grammars of diverse languages exhibit common features, there are two reasons to be suspicious of the inference to the doctrine of linguistic universals. The first, a point noted by Hilary Putnam,<sup>11</sup> is that even if all extant languages exhibit non-trivial common properties, there are explanations that could be offered for this fact which would not trace the common properties to features of the innate language acquisition mechanism. It might, for example, be the case that all extant languages were descended from a single common ur-language whose details have been lost in pre-history. If so, then we might expect to find common properties in extant languages, though these properties need not be exhibited by all humanly possible languages. Thus the extant languages might prove to be a relatively uncharacteristic sample of the humanly possible languages, exhibiting only a small part of the possible diversity in humanly possible languages. The second reason to be suspicious of the inference from common properties in extant languages to the doctrine of linguistic universals turns on how grammars are verified. Like other empirical theories, grammars are massively undetermined by the evidence we can garner for them. Given one grammar that can handle a body of data, it is no trick to produce many others that can handle the same data equally well. In an effort to reduce this empirical underdetermination, grammarians adopt the obvious strategy of opting for that grammar which, among all those that are compatible with the data, most closely resembles the best grammars we have for other languages. Plainly, however, this strategy has great potential for generating specious universals, since if we *can* write a grammar exhibiting a feature common to grammars for other languages, then we *will* do so, even if the data available for the language at hand cannot distinguish between the grammar we write and others that do not exhibit the 'universal' feature.<sup>12</sup>

There is, of course, another and quite different strategy that might be adopted in attempting to confirm the existence of linguistic universals. To show that a feature is universal to humanly possible languages we need only demonstrate that languages failing to exhibit the feature are not learnable as first languages by human children. But if this is the best test of the universals hypotheses, then it is to be hoped that the best test will never be performed.

*University of Michigan*

## NOTES

<sup>1</sup> See, for example, [Aspects] Ch. 1, [Language] Ch. 3, [Reflections] Chs. 1 and 4.

<sup>2</sup> Cf. my [WESK], [GPI] and [G&I], also Schwartz [Knowing], and Cooper [Knowledge]. For my current view on the topic of 'tacit knowledge' cf. [Beliefs].

<sup>3</sup> Cf. Chomsky's [Reflections] pp. 198ff. and the references given there.

<sup>4</sup> This talk of 'Empiricist' theories of mind and learning rests on the assumption that there are interesting family resemblances among Classical Empiricist theories of mind and modern learning theories in the behaviorist tradition. For example, both Classical Empiricism and recent behaviorist learning theory characteristically attribute to the mind an innate learning mechanism which is relatively simple, imposes relatively little structure on the output of the learning process, and is neither species nor task specific. Similarly, there are important family resemblances among Classical Rationalist theories of mind and the sort of learning theory advocated by Chomsky and his followers. For more on this theme, see Chomsky [Cartesian], [Language] ch. 1, and [Reflections] pp. 217ff.

<sup>5</sup> Perhaps a word is in order on the 'bit more'. The investigator presumably brings with him some cognitive mechanism that enables him to *think up* hypotheses. Empiricist account of mind, however, are generally quite uninformative about the process of hypothesis formation. Beyond the strategy of generalization and extrapolation, they have few suggestions on how we actually come up with an hypothesis to test. But this leaves the Empiricist with no substantive theory about how we discover the vast majority of serious scientific hypotheses. Often enough Empiricist theorists will be quite open about the deficiency, relegating the problem to the psychology of imagination or some similar domain.

<sup>6</sup> Let me add important caveat. Certain type II-R theories might actually have relatively modest needs for innate storage. Thus suppose all humanly possible languages have grammars which are identical save for a few details. — The analogy is the library specializing in books which differ only in the final chapter. — If this were true, then it would be relatively easy to store tens of thousands of grammars, since the redundant information need only be recorded once.

<sup>7</sup> Although paradigm cases of type II theories will be readily distinguishable from paradigm cases of type III theories, the boundary between the two types of theories is not all that sharp. Many sorts of acquisition theories might be constructed which straddle the divide between the two types. As an example, consider the type II-R theory outlined in Note 6. Here we have a sort of primitive production system which specifies the core of rules shared by all the grammars, as in the Library Model, then lists a number of different variations each of which, appended to the core rules, yields a complete grammar.

<sup>8</sup> I.e.  $G$  may be rewritten as  $G_1$ , or as  $G_2$ , ..., or as  $G_{17}$ .

<sup>9</sup> For an interesting discussion of some putative examples of universals along with a critique of the 'evidence' for them, see Cooper [Knowledge] Chapter 9.

<sup>10</sup> Chomsky has invoked instances of this argument on many occasions. Here is an example from his recent *Reflections on Language* (Pantheon Books, New York, 1975). Chomsky begins with a definition of 'universal grammar'. "Let us define 'universal grammar' (UG) as the system of principles, conditions, and rules that are elements or properties of all human languages not merely by accident but by necessity — of course, I mean biological, not logical, necessity. Thus UG can be taken as expressing 'the essence of human language'. UG will be invariant among humans. UG will specify what language learning must achieve, if it takes place successfully. Thus UG will be a significant component of LT (H, L), ['LT (H, L)'] has previously been defined as (p. 28) 'the system of mechanisms and principles put to work in acquisition of knowledge of language'.] What is learned, the cognitive structure attained, must have the properties of UG, though it will have other properties as well, accidental properties. If we were to construct a

language violating UG, we would find that it could not be learned by LG (H, L). That is, it would not be learnable under normal conditions of access and exposure to data.” (p. 29) A page later Chomsky tells us: “We can gain some insight into UG, hence LT (H, L), whenever we find properties of language that can be reasonably supposed not to have been learned.” (p. 30) He then spends several pages establishing that in processing declarative sentences an English speaker invokes a ‘structure-dependent’ rule. The next step is to ask why. “... The scientist must ask why it is that the child unerringly makes use of the structure-dependent rule postulated in hypothesis 2, rather than the simpler structure-independent rule of hypothesis 1. There seems to be no explanation in terms of ‘communicative efficiency’ or similar considerations. It is certainly absurd to argue that children are trained to use the structure-dependent rule, in this case. In fact, the problem never arises in language learning. A person may go through a considerable part of his life without ever facing relevant evidence, but he will have no hesitation in using the structure-dependent rule, even if all of his experience is consistent with hypothesis 1. The only reasonable conclusion is that UG contains the principle that all such rules must be structure-dependent. That is, the child’s mind (specifically its component LT (H, L)) contains the instruction: construct a structure-dependent rule, ignoring all structure independent rules. The principle of structure-dependence is not learned, but forms part of the conditions for language learning.” (pp. 32–3) The final step in Chomsky’s argument is to infer the universality of structure-dependence. “To corroborate this conclusion about UG (hence about LT (H, L)), the scientist will ask whether other rules of English are invariably structure-dependent. So far as we know the answer is positive. ... Having gotten this far, the scientist will conclude that other languages must have the same property, on the assumption that humans are not specifically designed to learn one rather than another language, say English rather than Japanese. On this reasonable assumption, the principle of structure-dependence... must hold universally if it holds for English.” (p. 33)

<sup>11</sup> In [Innateness].

<sup>12</sup> For more on this theme, cf. my [GPI] and [C&I].

#### BIBLIOGRAPHY

- Chomsky, Noam [Aspects], *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Mass., 1965.
- Chomsky, Noam [Cartesian], *Cartesian Linguistics*, Harper & Row, New York, 1966.
- Chomsky, Noam [Language], *Language and Mind*, Harcourt, Brace and World, New York, 1968.
- Chomsky, Noam [Reflections], *Reflections on Language*, Pantheon Books, New York 1975.
- Cooper, David [Knowledge], *Knowledge of Language*, Prism Press, London, and Humanities Press, New York, 1975.
- Putnam, Hilary [Innateness], ‘The “Innateness Hypothesis” and Explanatory Models in Linguistics’, *Synthese* 17 (1967). Reprinted in Stich [Innate].
- Schwartz, Robert [Knowing], ‘On Knowing a Grammar’, in S. Hook (ed.), *Language and Philosophy*, NYU Press, New York, 1969.
- Stich, Stephen P. [WESK], ‘What Every Speaker Knows’, *Philosophical Review* 80 (1971).
- Stich, Stephen P. [GPI], ‘Grammar, Psychology and Indeterminacy’, *Journal of Philosophy* 69 (1972).
- Stich, Stephen P. [C&I], ‘Competence and Indeterminacy’, in David Cohen and Jesica R. Wirth (eds.), *Testing Linguistic Hypotheses*, Hemisphere Publishing-John Wiley & Sons, Washington, D.C., 1975.

Stich, Stephen P. [Beliefs], 'Beliefs and Sub-Doxastic States'.

Stich, Stephen P. [Innate], *Innate Ideas*, University of California Press, Berkeley, 1975.