

NOAM CHOMSKY

RECENT CONTRIBUTIONS  
TO THE THEORY OF INNATE IDEAS

*Summary of Oral Presentation*

I think that it will be useful to separate two issues in the discussion of our present topic – one is the issue of historical interpretation, namely, what in fact was the content of the classical doctrine of innate ideas, let us say, in Descartes and Leibniz; the second is the substantive issue, namely, in the light of the information presently available, what can we say about the prerequisites for the acquisition of knowledge – what can we postulate regarding the psychologically a priori principles that determine the character of learning and the nature of what is acquired.

These are independent issues; each is interesting in its own right, and I will have a few things to say about each. What I would like to suggest is that contemporary research supports a theory of psychological a priori principles that bears a striking resemblance to the classical doctrine of innate ideas. The separateness of these issues must, nevertheless, be kept clearly in mind.

The particular aspect of the substantive issue that I will be concerned with is the problem of acquisition of language. I think that a consideration of the nature of linguistic structure can shed some light on certain classical questions concerning the origin of ideas.

To provide a framework for the discussion, let us consider the problem of designing a model of language-acquisition, an abstract 'language acquisition device' that duplicates certain aspects of the achievement of the human who succeeds in acquiring linguistic competence. We can take this device to be an input-output system

data → LA → knowledge

To study the substantive issue, we first attempt to determine the nature of the output in many cases, and then to determine the character of the function relating input to output. Notice that this is an entirely empirical matter; there is no place for any dogmatic or arbitrary assumptions about the intrinsic, innate structure of the device LA. The problem is quite analogous to the problem of studying the innate principles that make it

possible for a bird to acquire the knowledge that expresses itself in nest-building or in song-production. On a priori grounds, there is no way to determine the extent to which an instinctual component enters into these acts. To study this question, we would try to determine from the behavior of the mature animal just what is the nature of its competence, and we would then try to construct a second-order hypothesis as to the innate principles that provide this competence on the basis of presented data. We might deepen the investigation by manipulating input conditions, thus extending the information bearing on this input-output relation. Similarly, in the case of language-acquisition, we can carry out the analogous study of language-acquisition under a variety of different input conditions, for example, with data drawn from a variety of languages.

In either case, once we have developed some insight into the nature of the resulting competence, we can turn to the investigation of the innate mental functions that provide for the acquisition of this competence. Notice that the conditions of the problem provide an upper bound and a lower bound on the structure that we may suppose to be innate to the acquisition device. The upper bound is provided by the diversity of resulting competence – in our case, the diversity of languages. We cannot impose so much structure on the device that acquisition of some attested language is ruled out. Thus we cannot suppose that the specific rules of English are innate to the device and these alone, since this would be inconsistent with the observation that Chinese can be learned as readily as English. On the other hand, we must attribute to the device a sufficiently rich structure so that the output can be attained within the observed limits of time, data, and access.

To repeat, there is no reason for any dogmatic assumptions about the nature of LA. The only conditions we must meet in developing such a model of innate mental capacity are those provided by the diversity of language, and by the necessity to provide empirically attested competence within the observed empirical conditions.

When we face the problem of developing such a model in a serious way, it becomes immediately apparent that it is no easy matter to formulate a hypothesis about innate structure that is rich enough to meet the condition of empirical adequacy. The competence of an adult, or even a young child, is such that we must attribute to him a knowledge of language that extends far beyond anything that he has learned. Compared with the

number of sentences that a child can produce or interpret with ease, the number of seconds in a lifetime is ridiculously small. Hence the data available as input is only a minute sample of the linguistic material that has been thoroughly mastered, as indicated by actual performance. Furthermore, great diversity of input conditions does not lead to a wide diversity in resulting competence, so far as we can detect. Furthermore, vast differences in intelligence have only a small effect on resulting competence. We observe further that the tremendous intellectual accomplishment of language acquisition is carried out at a period of life when the child is capable of little else, and that this task is entirely beyond the capacities of an otherwise intelligent ape. [Such observations as these lead one to suspect, from the start, that we are dealing with a species-specific capacity with a largely innate component.] It seems to me that this initial expectation is strongly supported by a deeper study of linguistic competence. There are several aspects of normal linguistic competence that are crucial to this discussion.

#### I. CREATIVE ASPECT OF LANGUAGE USE

By this phrase I refer to the ability to produce and interpret new sentences in independence from 'stimulus control' – i.e., external stimuli or independently identifiable internal states. The normal use of language is 'creative' in this sense, as was widely noted in traditional rationalist linguistic theory. The sentences used in everyday discourse are not 'familiar sentences' or 'generalizations of familiar sentences' in terms of any known process of generalization. [In fact, even to speak of 'familiar sentences' is an absurdity.] The idea that sentences or sentence-forms are learned by association or conditioning or 'training' as proposed in recent behaviorist speculations, is entirely at variance with obvious fact. More generally, it is important to realize that in no technical sense of these words can language use be regarded as a matter of 'habit' or can language be regarded as 'a complex of dispositions to respond'.

A person's competence can be represented by a *grammar*, which is a system of rules for pairing semantic and phonetic interpretations. Evidently, these rules operate over an infinite range. [Once a person has mastered the rules (unconsciously, of course), he is capable, in principle, of using them to assign semantic interpretations to signals quite inde-

## RECENT CONTRIBUTIONS TO THE THEORY OF INNATE IDEAS

pendently of whether he has been exposed to them or their parts, as long as they consist of elementary units that he knows and are composed by the rules he has internalized. The central problem in designing a language acquisition device is to show how such a system of rules can emerge, given the data to which the child is exposed. In order to gain some insight into this question, one naturally turns to a deeper investigation of the nature of grammars. I think real progress has been made in recent years in our understanding of the nature of grammatical rules and the manner in which they function to assign semantic interpretations to phonetically represented signals, and that it is precisely in this area that one can find results that have some bearing on the nature of a language-acquisition device.

### II. ABSTRACTNESS OF PRINCIPLES OF SENTENCE INTERPRETATION

A grammar consists of syntactic rules that generate certain underlying abstract objects, and rules of semantic and phonological interpretation that assign an intrinsic meaning and an ideal phonetic representation to these abstract objects.

Concretely, consider the sentence 'The doctor examined John'. The phonetic form of this sentence depends on the intrinsic phonological character of its minimal items ('The', 'doctor', 'examine', 'past tense', 'John'), the bracketing of the sentence (that is, as [[[the] [doctor]] [[examined] [John]]]), and the categories to which the bracketed elements belong (that is, the categories 'Sentence', 'Noun-Phrase', 'Verb-Phrase', 'Verb', 'Noun', 'Determiner', in this case). We can define the 'surface structure' of an utterance as its labeled bracketing, where the brackets are assigned appropriate categorial labels from a fixed, universal set. It is transparent that grammatical relations (e.g., 'Subject-of', 'Object-of', etc.) can be defined in terms of such a labeled bracketing. With terms defined in this way, we can assert that there is very strong evidence that the phonetic form of a sentence is determined by its labeled bracketing by phonological rules that operate in accordance with certain very abstract but quite universal principles of ordering and organization.

The meaning of the sentence 'the doctor examined John' is, evidently, determined from the meanings of its minimal items by certain general rules that make use of the grammatical relations expressed by the labeled

bracketing. Let us define the 'deep structure' of a sentence to be that labeled bracketing that determines its intrinsic meaning, by application of these rules of semantic interpretation. In the example just given, we would not be far wrong if we took the deep structure to be identical with the surface structure. But it is obvious that these cannot in general be identified. Thus consider the slightly more complex sentences: 'John was examined by the doctor'; 'someone persuaded the doctor to examine John'; 'the doctor was persuaded to examine John'; 'John was persuaded to be examined by the doctor'. Evidently, the grammatical relations among *doctor*, *examine*, and *John*, as expressed by the deep structure, must be the same in all of these examples as the relations in 'the doctor examined John'. But the surface structures will differ greatly.

Furthermore, consider the two sentences:

someone expected the doctor to examine John  
 someone persuaded the doctor to examine John.

It is clear, in this case, that the similarity of surface structure masks a significant difference in deep structure, as we can see, immediately, by replacing 'the doctor to examine John' by 'John to be examined by the doctor' in the two cases.

So far, I have only made a negative point, namely, that deep structure is distinct from surface structure. Much more important is the fact that there is very strong evidence for a particular solution to the problem of how deep and surface structures are related, and how deep and surface structures are formed by the syntactic component of the grammar. The details of this theory need not concern us for the present. A crucial feature of it, and one which seems inescapable, is that it involves formal manipulations of structures that are highly abstract, in the sense that their relation to signals is defined by a long sequence of formal rules, and that, consequently, they have nothing remotely like a point by point correspondence to signals. Thus sentences may have very similar underlying structures despite great diversity of physical form, and diverse underlying structures despite similarity of surface form. A theory of language acquisition must explain how this knowledge of abstract underlying forms and the principles that manipulate them comes to be acquired and freely used.

## RECENT CONTRIBUTIONS TO THE THEORY OF INNATE IDEAS

### III. UNIVERSAL CHARACTER OF LINGUISTIC STRUCTURE

So far as evidence is available, it seems that very heavy conditions on the form of grammar are universal. Deep structures seem to be very similar from language to language, and the rules that manipulate and interpret them also seem to be drawn from a very narrow class of conceivable formal operations. There is no a priori necessity for a language to be organized in this highly specific and most peculiar way. There is no sense of 'simplicity' in which this design for language can be intelligibly described as 'most simple'. Nor is there any content to the claim that this design is somehow 'logical'. Furthermore, it would be quite impossible to argue that this structure is simply an accidental consequence of 'common descent'. Quite apart from questions of historical accuracy, it is enough to point out that this structure must be rediscovered by each child who learns the language. The problem is, precisely, to determine how the child determines that the structure of his language has the specific characteristics that empirical investigation of language leads us to postulate, given the meagre evidence available to him. Notice, incidentally, that the evidence is not only meagre in scope, but very degenerate in quality. Thus the child learns the principles of sentence formation and sentence interpretation on the basis of a corpus of data that consists, in large measure, of sentences that deviate in form from the idealized structures defined by the grammar that he develops.

Let us now return to the problem of designing a language acquisition device. The available evidence shows that the output of this device is a system of recursive rules that provide the basis for the creative aspect of language use and that manipulate highly abstract structures. Furthermore, the underlying abstract structures and the rules that apply to them have highly restricted properties that seem to be uniform over languages and over different individuals speaking the same language, and that seem to be largely invariant with respect to intelligence and specific experience. An engineer faced with the problem of designing a device meeting the given input-output conditions would naturally conclude that the basic properties of the output are a consequence of the design of the device. Nor is there any plausible alternative to this assumption, so far as I can see. More specifically, we are led by such evidence as I have mentioned to suppose that this device in some manner incorporates: a phonetic

theory that defines the class of possible phonetic representations; a semantic theory that defines the class of possible semantic representations; a schema that defines the class of possible grammars; a general method for interpreting grammars that assigns a semantic and phonetic interpretation to each sentence, given a grammar; a method of evaluation that assigns some measure of 'complexity' to grammars.

Given such a specification, the device might proceed to acquire knowledge of a language in the following way: the given schema for grammar specifies the class of possible hypotheses; the method of interpretation permits each hypothesis to be tested against the input data; the evaluation measure selects the highest valued grammar compatible with the data. Once a hypothesis – a particular grammar – is selected, the learner knows the language defined by this grammar; in particular, he is capable of pairing semantic and phonetic interpretations over an indefinite range of sentences to which he has never been exposed. Thus his knowledge extends far beyond his experience and is not a 'generalization' from his experience in any significant sense of 'generalization' (except, trivially, the sense defined by the intrinsic structure of the language acquisition device).

Proceeding in this way, one can seek a hypothesis concerning language acquisition that falls between the upper and lower bounds, discussed above, that are set by the nature of the problem. Evidently, for language learning to take place the class of possible hypotheses – the schema for grammar – must be heavily restricted.

This account is schematic and idealized. We can give it content by specifying the language acquisition system along the lines just outlined. I think that very plausible and concrete specifications can be given, along these lines, but this is not the place to pursue this matter, which has been elaborately discussed in many publications on transformational generative grammar.

I have so far been discussing only the substantive issue of the prerequisites for acquisition of knowledge of language, the a priori principles that determine how and in what form such knowledge is acquired. Let me now try to place this discussion in its historical context.

First, I mentioned three crucial aspects of linguistic competence: (1) creative aspect of language use; (2) abstract nature of deep structure; (3) apparent universality of the extremely special system of mechanisms formalized now as transformational grammar. It is interesting to observe

that these three aspects of language are discussed in the rationalist philosophy of the 17th century and its aftermath, and that the linguistic theories that were developed within the framework of this discussion are, in essence, theories of transformational grammar.

Consequently, it would be historically accurate to describe the views regarding language structure just outlined as a rationalist conception of the nature of language. Furthermore, I employed it, again, in the classical fashion, to support what might fairly be called a rationalist conception of acquisition of knowledge, if we take the essence of this view to be that the general character of knowledge, the categories in which it is expressed or internally represented, and the basic principles that underlie it, are determined by the nature of the mind. In our case, the schematism assigned as an innate property to the language acquisition device determines the form of knowledge (in one of the many traditional senses of 'form'). The role of experience is only to cause the innate schematism to be activated, and then to be differentiated and specified in a particular manner.

In sharp contrast to the rationalist view, we have the classical empiricist assumption that what is innate is (1) certain elementary mechanisms of peripheral processing (a receptor system), and (2) certain analytical mechanisms or inductive principles or mechanisms of association. What is assumed is that a preliminary analysis of experience is provided by the peripheral processing mechanisms and that one's concepts and knowledge, beyond this, are acquired by application of the innate inductive principles to this initially analyzed experience. Thus only the procedures and mechanisms for acquisition of knowledge constitute an innate property. In the case of language acquisition, there has been much empiricist speculation about what these mechanisms may be, but the only relatively clear attempt to work out some specific account of them is in modern structural linguistics, which has attempted to elaborate a system of inductive analytic procedures of segmentation and classification that can be applied to data to determine a grammar. It is conceivable that these methods might be somehow refined to the point where they can provide the surface structures of many utterances. It is quite inconceivable that they can be developed to the point where they can provide deep structures or the abstract principles that generate deep structures and relate them to surface structures. This is not a matter of further refinement, but of an entirely different approach to the question. Similarly, it is difficult to imagine how



the vague suggestions about conditioning and associative nets that one finds in philosophical and psychological speculations of an empiricist cast might be refined or elaborated so as to provide for attested competence. A system of rules for generating deep structures and relating them to surface structures, in the manner characteristic of natural language, simply does not have the properties of an associative net or a habit family; hence no elaboration of principles for developing such structures can be appropriate to the problem of designing a language acquisition device.

I have said nothing explicit so far about the doctrine that there are innate ideas and innate principles of various kinds that determine the character of what can be known in what may be a rather restricted and highly organized way. In the traditional view a condition for these innate mechanisms to become activated is that appropriate stimulation must be presented. This stimulation provides the occasion for the mind to apply certain innate interpretive principles, certain concepts that proceed from 'the power of understanding' itself, from the faculty of thinking rather than from external objects. To take a typical example from Descartes (Reply to Objections, V): "...When first in infancy we see a triangular figure depicted on paper, this figure cannot show us how a real triangle ought to be conceived, in the way in which geometers consider it, because the true triangle is contained in this figure, just as the statue of Mercury is contained in a rough block of wood. But because we already possess within us the idea of a true triangle, and it can be more easily conceived by our mind than the more complex figure of the triangle drawn on paper, we, therefore, when we see the composite figure, apprehend not it itself, but rather the authentic triangle" (Haldane and Ross, vol. II, p. 227). In this sense, the idea of triangle is innate. For Leibniz what is innate is certain principles (in general, unconscious), that "enter into our thoughts, of which they form the soul and the connection". "Ideas and truths are for us innate as inclinations, dispositions, habits, or natural potentialities." Experience serves to elicit, not to form, these innate structures. Similar views are elaborated at length in rationalist speculative psychology.

It seems to me that the conclusions regarding the nature of language acquisition, discussed above, are fully in accord with the doctrine of innate ideas, so understood, and can be regarded as providing a kind of substantiation and further development of this doctrine. Of course,

such a proposal raises nontrivial questions of historical interpretation.

What does seem to me fairly clear is that the present situation with regard to the study of language learning, and other aspects of human intellectual achievement of comparable intricacy, is essentially this. We have a certain amount of evidence about the grammars that must be the output of an acquisition model. This evidence shows clearly that know- 1.  
ledge of language cannot arise by application of step-by-step inductive 2.  
operations (segmentation, classification, substitution procedures, 'analogy', association, conditioning, and so on) of any sort that have been developed or discussed within linguistics, psychology, or philosophy. Further empiricist speculations contribute nothing that even faintly suggests a way of overcoming the intrinsic limitations of the methods that have so far been proposed and elaborated. Furthermore, there are no other grounds for pursuing these empiricist speculations, and avoiding what would be the normal assumption, unprejudiced by doctrine, that one would formulate if confronted with empirical evidence of the sort sketched above. There is, in particular, nothing known in psychology or physiology that suggests that the empiricist approach is well-motivated, or that gives any grounds for skepticism concerning the rationalist alternative sketched above.

For further discussion of the question of historical interpretation, see Chomsky, *Aspects of the Theory of Syntax* (1965), ch. 1, and *Cartesian Linguistics* (1966). For further discussion of matters touched on here, see also Chomsky, 'Explanatory Models in Linguistics', in *Logic, Methodology and Philosophy of Science*, ed. by E. Nagel, P. Suppes, and A. Tarski (1962); J. Katz, *The Philosophy of Language* (1966); P. M. Postal, Review of A. Martinet, *Elements of General Linguistics* (forthcoming); and the selections in section VI of *The Structure of Language, Readings in the Philosophy of Language*, ed. by J. Fodor and J. Katz (1964).

*Massachusetts Institute of Technology*