HILARY PUTNAM

THE 'INNATENESS HYPOTHESIS' AND EXPLANATORY MODELS IN LINGUISTICS

I. THE INNATENESS HYPOTHESIS

The 'innateness hypothesis' (henceforth, the 'I.H.') is a daring – or apparently daring; it may be meaningless, in which case it is not daring – hypothesis proposed by Noam Chomsky. I owe a debt of gratitude to Chomsky for having repeatedly exposed me to the I.H.; I have relied heavily in what follows on oral communications from him; and I beg his pardon in advance if I misstate the I.H. in any detail, or misrepresent any of the arguments for it. In addition to relying upon oral communications from Chomsky, I have also relied upon Chomsky's paper 'Explanatory Models in Linguistics', in which the I.H. plays a considerable rôle.

To begin, then, the I.H. is the hypothesis that the human brain is 'programmed' at birth in some quite *specific* and *structured* aspects of human natural language. The details of this programming are spelled out in some detail in 'Explanatory Models in Linguistics'. We should assume that the speaker has 'built in'[1] a function which assigns weights to the grammars $G_1, G_2, G_3, \ldots$ in a certain class Σ of transformational grammars. Σ is not the class of all *possible* transformational grammars; rather all the members of Σ have some quite strong similarities. These similarities appear as 'linguistic universals' – i.e., as characteristics of *all* human natural languages. If intelligent non-terrestrial life – say, Martians – exists, and if the 'Martians' speak a language whose grammar does not belong to the subclass Σ of the class of all transformational grammars, then, I have heard Chomsky maintain, humans (except possibly for a few geniuses or linguistic experts) would be unable to learn Martian; a human child brought up by Martians would fail to acquire language; and Martians would, conversely, experience similar difficulties with human tongues. (Possible difficulties in *pronunciation* are not at issue here, and may be assumed *not* to exist for the purposes of this argument.) As examples of the similarities that all grammars of the subclass Σ are thought to possess (above the level of phonetics), we may mention the *active–*

12

passive distinction, the existence of a *non-phrase-structure* portion of the grammar, the presence of such major categories as *concrete noun, verb taking an abstract subject,* etc. The project of delimiting the class Σ may also be described as the project of defining a *normal form for grammars.* Conversely, according to Chomsky, any non-trivial normal form for grammars, such that correct and perspicuous grammars of all human languages can and should be written in that normal form, "constitutes, in effect, a hypothesis concerning the innate intellectual equipment of the child".[2]

Given such a highly *restricted* class Σ of grammars (highly restricted in the sense that grammars not in the class are perfectly conceivable, not more 'complicated' in any absolute sense than grammars in the class, and may well be employed by non-human speakers, if such there be), the performance of the human child in learning his native language may be understood as follows, according to Chomsky. He may be thought of as operating on the following 'inputs'[3]: a list of utterances, containing both grammatical and ungrammatical sentences; a list of corrections, which enable him to classify the input utterances *as* grammatical or ungrammatical; and some information concerning which utterances count as *repetitions* of earlier utterances. Simplifying slightly, we may say that, on this model, the child is supplied with a list of grammatical sentence *types* and a list of ungrammatical sentence *types.* He then 'selects' the grammar in Σ compatible with this information to which his weighting function assigns the highest weight. On this scheme, the general *form* of grammar is not learned from experience, but is 'innate', and the 'plausibility ordering' of grammars compatible with given data of the kinds mentioned is likewise 'innate'.

So much for a statement of the I.H. If I have left the I.H. vague at many points, I believe that this is no accident – for the I.H. seems to me to be *essentially* and *irreparably* vague – but this much of a statement may serve to indicate *what* belief it is that I stigmatize as irreparably vague.

A couple of remarks may suffice to give some idea of the rôle that I.H. is supposed to play in linguistics. Linguistics relies heavily, according to Chomsky, upon 'intuitions' of grammaticality. But *what* is an intuition of 'grammaticality' an intuition *of*? According to Chomsky, the sort of theory-construction programmatically outlined above is what is needed to give this question the only answer it can have or deserves to have. Presumable, then, to 'intuit' (or assert, or conjecture, etc.) that a sentence is grammatical is to 'intuit' (or assert, or conjecture, etc.) that the sentence

is generated by the highest-valued G_i in the class Σ which is such that it generates all the grammatical sentence types with which we have been supplied by the 'input' and none of the ungrammatical sentence types listed in the 'input'.[4]

Chomsky also says that the G_i which receives the highest value must do *more* than agree with 'intuitions' of grammaticality; it must account for certain ambiguities, for example.[5] At the same time, unfortunately, he lists no semantical information in the input, and he conjectures[6] that a child needs semantical information only to "provide motivation for language learning", and not to arrive at the *formal* grammar of its language. Apparently, then, the fact that a grammar which agrees with a sufficient amount of 'input' must be in the class Σ to be 'selected' by the child is what rules out grammars that generate all and only the grammatical sentences of a given natural language, but fail to correctly 'predict'[7] ambiguities (cf. E. M. in L., p. 533).

In addition to making clear what it *is* to be grammatical, Chomsky believes that the I.H. confronts the linguist with the following tasks: To *define* the normal form for grammars described above, and to *define* the weighting function. In *Syntactic Structures* Chomsky, indeed, gives this as an objective for linguistic theory: to give an *effective* procedure for choosing between rival grammars.

Lastly, the I.H. is supposed to justify the claim that what the linguist provides is "a hypothesis about the innate intellectual equipment that a child brings to bear in language learning".[8] Of course, even if language is *wholly* learned, it is still true that linguistics "characterizes the linguistic abilities of the nature speaker"[9], and that a grammar "could properly be called an explanatory model of the linguistic intuition of the native speaker".[10] However, one could with equal truth say that a driver's manual "characterizes the car-driving abilities of the mature driver" and that a calculus text provides "an explanatory model of the calculus-intuitions of the mathematician". Clearly, it is the idea that *these* abilities and *these* intuitions are close to the human *essence*, so to speak, that gives linguistics its 'sex appeal', for Chomsky at least.

II. THE SUPPOSED EVIDENCE FOR THE I.H.

A number of empirical facts and alleged empirical facts have been

14

advanced to support the I.H. Since limitations of space make it impossible to describe all of them here, a few examples will have to suffice.

(a) The *ease* of the child's original language learning "A young child is able to gain perfect mastery of a language with incomparably greater ease [*than an adult* – H.P.] and without any explicit instruction. Mere exposure to the language, and for a remarkably short period, seems to be all that the normal child requires to develop the competence of the native speaker".[11]

(b) The fact that reinforcement, "in any interesting sense", seems to be unnecessary for language learning. Some children have apparently even learned to speak without *talking*[12], and then displayed this ability at a relatively late age to startled adults who had given them up for mutes.

(c) The ability to "develop the competence of the native speaker" has been said not to depend on the intelligence level. Even quite low I.Q.'s 'internalize' the grammar of their native language.

(d) The 'linguistic universals' mentioned in the previous section are allegedly accounted for by the I.H.

(e) Lastly, of course, there is the 'argument' that runs "*what else* could account for language learning?" The task is so incredibly complex (analogous to learning, at least implicitly, a complicated physical theory, it is said), that it would be miraculous if even one tenth of the human race accomplished it without 'innate' assistance. (This is like Marx's 'proof' of the Labour Theory of Value in *Capital*, vol. III, which runs, in essence, "*What else* could account for the fact that commodities have different value *except* the fact that the labor-content is different?".)

III. CRITICISM OF THE ALLEGED EVIDENCE

A. The Irrelevance of Linguistic Universals

1. Not surprising on any theory

Let us consider just how surprising the 'linguistic universals' cited above really are. Let us assume for the purpose a community of Martians whose 'innate intellectual equipment' may be supposed to be as different from the human as is compatible with their being able to speak a language at all. What could we expect to find in their language?

If the Martians' brains are not vastly richer than ours in complexity,

then they, like us, will find it possible to employ a practically infinite set of expressions only if those expressions possess a 'grammar' – i.e., if they are built up by recursive rules from a limited stock of basic forms. Those basic forms need not be built up out of a *short* list of phonemes – the Martians might have vastly greater memory capacity than we do – but if Martians, like humans, find rote learning difficult, it will not be surprising if they too have *short* lists of phonemes in their languages.

Are the foregoing reflections arguments *for* or *against* the I.H.? I find it difficult to tell. If belief in 'innate intellectual equipment' is *just* that, then how *could* the I.H. be false? How could something with *no* innate intellectual equipment *learn* anything? *To be sure,* human 'innate intellectual equipment' is relevant to language learning; if this means that such parameters as memory span and memory capacity play a crucial role. But what rank Behaviorists is supposed to have ever denied *this*? On the other hand, that a particular mighty arbitrary set Σ of grammars is 'built in' to the brain of *both* Martians and Humans is *not* a hypothesis we would have to invoke to account for *these* basic similarities.

But for what similarities above the level of phonetics, where constitutional factors play a large role for obvious reasons, *would* the I.H. have to be invoked *save* in the trivial sense that memory capacity, intelligence, needs, interests, etc., are all relevant to language learning, and all depend, in part, on the biological makeup of the organism? If Martians are such strange creatures that they have no interest in physical objects, for example, their language will contain no concrete nouns; but would not this be *more*, not *less* surprising, on any *reasonable* view, than their having an interest in physical objects? (Would it be surprising if Martian contained devices for forming truth-functions and for quantification?)

Two more detailed points are relevant here. Chomsky has pointed out that no natural language has a phrase structure grammar. But this too is not surprising. The sentence 'John and Jim came home quickly' is not generated by a phrase-structure rule, in Chomsky's formalization of English grammar. But the sentence 'John came home quickly and Jim came home quickly' *is* generated by a phrase-structure rule in the grammar of mathematical logic, and Chomsky's famous 'and-transformation' is just an abbreviation rule. Again, the sentence 'That was the lady I saw you with last night' is not generated by a phrase-structure rule in English, or at least not in Chomsky's description of English. But the sentence

'That is $\iota x(x$ is a lady and I saw you with x last night)' is generated by a phrase-structure rule in the grammar of mathematical logic. And again the idiomatic English sentence *can* be obtained from its phrase-structure counterpart by a simple rule of abbreviation. Is it really surprising, does it really point to anything more interesting than *general intelligence*, that these operations which break the bounds of phrase-structure grammar appear in every natural language?[13]

Again, it may appear startling at first blush that such categories as noun, verb, adverb, etc. have 'universal' application. But, as Curry has pointed out, it is too easy to multiply 'facts' here. If a language contains nouns – that is, a phrase-structure category which contains the proper names – it contains noun phrases, that is, phrases which occupy the environments of nouns. If it contains noun phrases it contains verb phrases – phrases which when combined with a noun phrase by a suitable construction yield sentences. If it contains verb phrases, it contains adverb phrases – phrases which, when combined with a verb phrase yield a verb phrase. Similarly, adjective phrases, etc., can be defined in terms of the *two* basic categories 'noun' and 'sentence'. Thus the existence of nouns is all that has to be explained. And this reduces to explaining two facts: (1) The fact that all natural languages have a large phrase structure portion in their grammar, in the sense just illustrated, in spite of the effect of what Chomsky calls 'transformations'. (2) The fact that all natural languages contain proper names. But (1) is not surprising in view of the fact that phrase-structure rules are extremely simple algorithms. Perhaps Chomsky would reply that 'simplicity' is subjective here, but this is just not so. The fact is that all the natural measures of complexity of an algorithm – size of the machine table, length of computations, time, and space required for the computation – lead to the same result here, quite independently of the detailed structure of the computing machine employed. Is it surprising that algorithms which are 'simplest' for virtually any computing system we can conceive of are also simplest for naturally evolved 'computing systems'? And (2) – the fact that all natural languages contain proper names – is not surprising in view of the utility of such names, and the difficulty of always finding a definite description which will suffice instead.

Once again, 'innate' factors are relevant *to be sure* – if choosing *simple* algorithms as the basis of the grammar is 'innate', and if the need for

identifying persons rests on something innate – but what Behaviorist would or should be surprised? Human brains are computing systems and subject to some of the constraints that effect all computing systems; human beings have a natural interest in one another. If *that* is 'innateness', well and good!

2. *Linguistic universals could be accounted for, even if surprising, without invoking the I.H.*

Suppose that language-using human beings evolved *independently* in two or more places. Then, if Chomsky were *right*, there should be two or more *types* of human beings descended from the two or more original populations, and normal children of each type should fail to learn the languages spoken by the other types. Since we do not observe this, since there is only *one* class Σ built into *all* human brains, we have to conclude (if the I.H. is true) that language-using is an evolutionary 'leap' that occurred only *once*. But in that case, it is overwhelmingly likely that all human languages are descended from a single original language, and that the existence today of what are called 'unrelated' languages is accounted for by the great lapse of time and by countless historical changes. This is, indeed, likely even if the I.H. is false, since the human race itself is now generally believed to have resulted from a single evolutionary 'leap', and since the human population was extremely small and concentrated for millenia, and only gradually spread from Asia to other continents. Thus, even if language using was learned or invented rather than 'built in', or even if only some general dispositions in the direction of language using are 'built in'[14], it is likely that some one group of humans first developed language as we know it, and then spread this through conquest or imitation to the rest of the human population. Indeed, we do know that this is just how *alphabetic* writing spread. In any case, I repeat, this hypothesis – a single origin for human language – is certainly *required* by the I.H., but much weaker than the I.H.

But just this *consequence* of the I.H. is, in fact, enough to account for 'linguistic universals'! For, if all human languages are descended from a common parent, then just such highly useful features of the common parent as the presence of some kind of quantifiers, proper names, nouns, and verbs, etc., would be expected to survive. Random variation may, indeed, alter many things; but that it should fail to strip language of

proper names, or common nouns, or quantifiers, is not *so* surprising as to require the I.H.

B. *The 'ease' of Language Learning is not clear*

Let us consider somewhat closely the 'ease' with which children do learn their native language. A typical 'mature' college student seriously studying a foreign language spends three hours a week in lectures. In fourteen weeks of term he is thus exposed to forty-two hours of the language. In four years he may pick up over 300 hours of the language, very little of which is actual listening to native informants. By contrast, direct method teachers estimate that 300 hours of direct-method teaching will enable one to converse fluently in a foreign language. Certainly 600 hours – say, 300 hours of direct-method teaching and 300 hours of reading – will enable any adult to speak and read a foreign language with ease, and to use an incomparably larger vocabulary than a young child.

It will be objected that the adult does not acquire a perfect accent. So what? The adult has been speaking one way all of his life, and has a huge set of habits to unlearn. What can equally well be accounted for by learning theory should not be cited as evidence for the I.H.

Now the child by the time it is four or five years old has been exposed to *vastly* more than 600 hours of direct-method instruction. Moreover, even if 'reinforcement' is not necessary, most children are consciously and repeatedly reinforced by adults in a host of ways – e.g., the constant repetition of simple one-word sentences ('cup', 'doggie') in the presence of babies. Indeed, any foreign adult living with the child for those years would have an incomparably better grasp of the language than the child does. The child indeed has a better accent. Also, the child's grammatical mistakes, which are numerous, arise not from carrying over previous language habits, but from not having fully acquired the first set. But it seems to me that this 'evidence' for the I.H. stands the facts on their head.

C. *Reinforcement another Issue*

As Chomsky is aware, the evidence is today slim that *any* learning requires reinforcement "in any interesting sense". Capablanca, for example, learned to play chess by simply watching adults play. This is comparable

to Macaulay's achievement in learning language without speaking. Non-geniuses normally do require practice both to speak correctly and to play chess. Yet probably anyone *could* learn to speak *or* to play chess without practice if muffled, in the first case, or not allowed to play, in the second case, with sufficiently prolonged observation.

D. Independence of Intelligence Level an Artifact

Every child learns to speak the native language. What does this mean? If it means that children do not make serious grammatical blunders, even by the standards of descriptive as opposed to prescriptive grammar, this is just not true for the young child. By nine or ten years of age this has ceased to happen, perhaps (I speak as a parent), but nine or ten years is enough time to become pretty darn good at *anything*. What is more serious is what 'grammar' *means* here. It does not include mastery of vocabulary, in which even many adults are deficient, nor ability to understand *complex* constructions, in which many adults are *also* deficient. It means purely and simply the ability to learn what every *normal* adult learns. Every normal adult learns what every normal adult learns. What this 'argument' reduces to is "Wow! How complicated a skill every normal adult learns. What else could it be but *innate*." Like the preceding argument, it reduces to the 'What Else' argument.

But what of the 'What Else?' argument? Just how impressed should we be by the failure of current learning theories to account for complex learning processes such as those involved in the learning of language? If Innateness were a *general* solution, perhaps we should be impressed. But the I.H. *cannot*, by its very nature, *be* generalized to handle all complex learning processes. Consider the following puzzle (called 'jump'):

```
        *   *   *
        *   *   *
*   *   *   *   *   *   *
*   *   *   ●   *   *   *
*   *   *   *   *   *   *
        *   *   *
        *   *   *
```
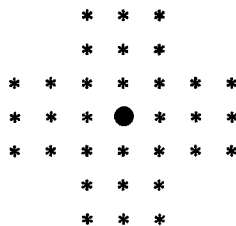
To begin with, all the holes but the center one are filled. The object of

the game is to remove all the pegs but one by 'jumping' (as in checkers) and to end with the one remaining peg in the center. A clever person can get the solution in perhaps eight or ten hours of experimentation. A not so clever person can get a 'near-solution' – two pegs left – in the same time. No program exists, to my knowledge, that would enable a computer to solve even the 'near solution' problem without running out of both time and space, even though the machine can spend the equivalent of many human lifetimes in experimentation. When we come to the discovery of even the simplest mathematical theorem the situation is even more striking. The theorems of mathematics, the solutions to puzzles, etc., cannot on *any* theory be *individually* 'innate'; what must be 'innate' are heuristics, i.e., learning strategies. In the absence of any knowledge of what *general multipurpose learning strategies* might even look like, the assertion that such strategies (which absolutely must exist and be employed by all humans) cannot account for this or that learning process, that the answer or an answer schema must be 'innate', is utterly unfounded.

I will be told, of course, that *everyone* learns his native language (as well as everyone does), and that not everyone solves puzzles or proves theorems. But everyone does learn pattern recognition, automobile driving, etc., and everyone in fact can solve many problems that no computer can solve. In conversation Chomsky has repeatedly used precisely such skills as these to support the idea that humans have an "innate conceptual space". Well and good, if true. *But that is no help. Let a complete 17th-century Oxford University education be innate if you like*; still the solution to 'jump' was not innate; the Prime Number Theorem was not innate; and so on. *Invoking 'Innateness' only postpones the problem of learning; it does not solve it.* Until we understand the strategies which make general learning possible – and vague talk of 'classes of hypotheses' – and 'weighting functions' is utterly useless here – no discussion of the *limits* of learning can even begin.

Massachusetts Institute of Technology

REFERENCES

[1] What 'built in' means is highly unclear in this context. The weighting function by itself determines only the relative ease with which various grammars can be learned by a

human being. If a grammar G_1 can be learned more easily than a grammar G_2, then doubtless this is 'innate' in the sense of being a fact about human learning *potential*, as opposed to a fact about what has been learned. But this sort of fact is what learning theory tries to account for; *not* the explanation being sought. It should be noticed that Chomsky has never offered even a schematic account of the sort of device that is supposed to be present in the brain, and that is supposed to do the job of selecting the highest weighted grammar compatible with the data. But only a description, or at least a theory, of such a device could properly be called an innateness *hypothesis* at all.

2 E. M. in L., p. 550.

3 E. M. in L., pp. 530–531.

4 I doubt that the child really is told which sentences it hears or utters are *ungrammatical*. At most it is told which are *deviant* – but it may not be told which are deviant for *syntactical* and which for *semantical* reasons.

5 Many of these – e.g., the alleged 'ambiguity' in 'the shooting of the elephants was heard' – *require coaching to detect*. The claim that grammar "explains the ability to recognize ambiguities" thus lacks the impressiveness that Chomsky believes it to have. I am grateful to Paul Ziff and Stephen Leeds for calling this point to my attention.

6 E. M. in L., p. 531, n. 5.

7 A grammar 'predicts' an ambiguity, in Chomsky's formalism, whenever it assigns two or more structural descriptions to the same sentence.

8 E. M. in L., p. 530.

9 E. M. in L., p. 530.

10 E. M. in L., p. 533.

11 E. M. in L., p. 529.

12 Macaulays *first* words, it is said, were: "Thank you, Madam, the agony has somewhat abated" (to a lady who had spilled hot tea on him).

13 Another example of a transformation is the 'active–passive' transformation (cf. *Syntactic Structures*). But (a) the presence of this, if it *is* a part of the grammar, is not surprising – why should not there be a systematic way of expressing the *converse* of a relation? – and (b) the argument for the existence of such a 'transformation' at all is extremely slim. It is contended that a grammar which 'defines' active and passive forms separately (this can be done by even a phrase-structure grammar) fails to represent something that every speaker knows, *viz.* that active and passive forms are *related*. But why must every *relation* be mirrored by *syntax*? Every 'speaker' of the canonical languages of mathematical logic is aware that each sentence $(x) (Fx \supset Gx)$ is related to a sentence $(x) (\bar{G}x \supset \bar{F}x)$; yet the definition of 'well formed formula' fails to mirror 'what every speaker knows' in this respect, and is not inadequate on that account.

14 It is very difficult to account for such phenomena as the spontaneous babbling of infants without *this* much 'innateness'. But this is not to say that a class Σ and a function f are 'built in', as required by the I.H.