REMARKS ON THE LANGUAGE UNIVERSALS RESEARCH II

Esa Itkonen

As the title indicates, this article is meant to be a sequel to Itkonen (1991b). In the previous article I claimed that the notion of *analogy*, if properly understood and defined, constitutes the cornerstone of the explanatory, typological-functional approach to language universals. I also criticized the way that the non-explanatory, innatist approach has attempted to dispose of analogy, especially in connection with the 'poverty of the stimulus' argument. In the present article I shall bring up additional reflections on these two approaches.[1]

1. THE GREENBERGIAN APPROACH

Greenberg (1966 [1963]) drew attention to correlations that pertain between various linguistic phenomena in the world's languages. Such correlations are typically expressed as *universal implications*: 'If a language has the property A, it has the property B'; or, more schematically, 'For all languages, if A, then B'. For instance: 'If a language has initial clusters, it has medial clusters'; 'If a language has gender categories in the noun, it has gender categories in the pronoun'; 'If a language has case-marking in intransitive subjects, it has case-marking in transitive subjects'.

Greenberg-type implicational universals became a centerpiece of the 'typological-functional school', and it goes without saying that they entailed genuine progress in linguistic theorizing. They made the notion of *general* linguistics look a little less like just an empty promise. And in particular, they were a useful antidote against the 'general' linguist's perverse inclination to study nothing but his native language. It seems, however, that Greenberg as well as several of his more prominent followers have misunderstood the nature of implicational universals to some extent. And the principal reason for this misunderstanding resides in the fact that they have understood less than perfectly the nature of the *sentence-type* employed to express the universals.

1.1. Implicational Universals and Causality

We have a natural tendency to give a *causal* interpretation to an implication 'If A, then B', in such a way that *A* and *B* stand, respectively, for the cause and the effect (cf. Wason & Johnson-Laird 1972, chs 7–8). Implications may be used just as well, however, to express a relation from effect to cause. Thus, the following sentences are equally possible: 'If it has been raining during the night, the streets are wet in the morning' (= cause-to-effect) and 'If the streets are wet in the morning, it has been raining during the night'(= effect-to-cause).[2] Notice that it is not possible to *predict* the (anterior) cause on the basis of the (posterior) effect. Notice also that while causes (with the associated laws) are standardly employed to *explain* their effects, it is not possible to do the opposite, i.e. to explain causes on the basis of their effects. (*'Why did it rain?' — 'Because the streets are wet.')[3]

If we look carefully at our examples of implicational universals, with the general structure 'For all languages, if A, then B', we notice that they express neither cause-to-effect relations nor effect-to-cause relations. Rather, they just express a pattern of *asymmetric co-occurrence*: if the less normal occurs, the normal occurs as well (but not vice versa). This pattern seems to be based on the general truth that primary needs are satisfied before secondary needs. Accordingly, our sentences might be reinterpreted as expressing the fact that if secondary needs are satisfied, then (it can be inferred or 'predicted' that) primary needs are, or have been, satisfied.[4] For instance, it is self-evident that the need for formal distinction is greater in the case of transitive subjects (which have to be distinguished both from the verbs and from the objects) than in the case of intransitive subjects (which have to be distinguished only from the verbs). The distribution of medial and initial clusters, in turn, rests on the need for the ease of articulation: if what is less easy to articulate occurs, what is more easy to articulate occurs too. The underlying idea is expressed e.g. by such a down-to-earth implication as 'If people have enough money for a new TV set, they have (or must have) enough money for food'.

Greenberg (1978b: 77) assumes that, in general, an implication like 'If A, then B' expresses a *causal* relation; and since *A* is 'dependent on' *B*

(1978a: 44), it follows that 'A' and 'B' should, in general, stand for the effect and the cause, respectively.[5] It should be clear, however, that the satisfaction of primary needs does not *cause* (although it makes possible) the satisfaction of secondary needs. The satisfaction of needs, whether primary or secondary, flows from a *common* (causal) source, with a certain order of priorities.

1.2. Implicational Universals and Deductive-Nomological Explanation

While Greenberg moves, as it were, from the consequent to the antecedent, Moravcsik (1978: 9) does the opposite. She assumes that since, in a sentence like 'If A, then B', *B* is *predicted* on the basis of *A*, it is also automatically *explained* by *A*.

This calls for comments. First, implicational universals do not permit any *genuine* predictions. It should be said that the presence of primary (or unmarked) cases is *inferred* (rather than 'predicted') from the presence of secondary (or marked) cases. Second, an inference of this type is not *explanatory* in any intuitively natural sense. The mere fact that, e.g., a language has overt case-marking for intransitive subjects does not *explain* in the least why it has overt case-marking for transitive subjects. If one wishes to give a genuine explanation, one has to refer to the greater or lesser need to make formal distinctions in the two cases. And one cannot do this without appealing, ultimately, to the notion of (unconscious) *rationality*. The interested reader can find a more detailed explanation in Itkonen (1983: 215–218).

It is possible, however, that the term 'explanation' is being deliberately used in an intuitively non-natural sense. This is evident from Hammond, Moravcsik, and Wirth (1988: 2). These authors repeat the claim that 'prediction', as employed in connection with implicational universals, equals explanation. What they seem to have in mind is the thesis of structural symmetry between (deductive-nomological) explanation and prediction (cf. Itkonen 1978: 1.2).

Hammond *et al.'s* (1988) argument may be rephrased as follows. A sentence like 'If a language has voiceless syllabic stops, it has voiced syllabic stops' expresses a 'general law' valid for all languages. If the 'antecedent condition' of some language L_n having voiceless syllabic stops is adduced next, then the ('particular') fact of L_n having voiced syllabic stops is taken to have been both predicted and explained by the combination of the law and of the antecedent condition. This is called an 'explanation', not because it makes the occurrence of voiced syllabic stops more comprehensible, for instance by revealing its causation, but because it — presumably — makes the occurrence of voiced syllabic stops *necessary*. This necessity in turn derives from the above-mentioned general law, because such laws by definition possess 'nomic necessity'.

The authors go on to give an even clearer illustration of what they mean by 'explanation'. Another general law is expressed by the unrestricted universal 'All languages have stops' (or more explicitly, 'For all x, if x is a language, then x has stops'). Once we have adduced the 'antecedent condition' that English is a language, we have presumably *explained* the fact that English has stops (namely by making it necessary).

It is quite true that these putative explanations conform to the schema of the Hempelian deductive-nomological explanation. But their very unnaturalness demonstrates that this schema is not an adequate explication of scientific explanation, mainly because it in itself does not guarantee that there is any reference to *causation*. This is indeed the general opinion in the modern philosophy of science (cf. Stegmüller 1974, chap. II, esp. pp. 191–199). For instance, the following inference does *not* (really) explain why the thing *a* is black:

For all *x*, if *x* is a crow, then *x* is black.
a is a crow
___
a is black

A genuine explanation would have to refer to the (causal, lawlike) mechanism that produces the black colour in the feathers of crows.

1.3. Implicational Universals and Natural Laws

The use of implicational universals as the major premisses of DN-explanations entails that they are considered as (analogous to) natural laws; and the Greenbergian approach rests on the assumption that implicational universals qualify as (expressions of) laws valid for *all* human languages (cf. Croft 1990: 48). The dubious nature of this assumption can be shown in the following way.

Let us assume that some biological characteristic divides an animal species *S* into three subspecies *S-1*, *S-2*, and *S-3*, and that all members of *S-1* have some property *A* (e.g., they die before reaching the age of 10), whereas the members of *S–2* and *S–3* do not have this property. Thus, the law that consists in having the constant property *A*, i.e. mortality-before-10, is valid for *S–1*, but not for *S*. Interestingly enough, this fact disappears from view, if we formulate it as a universal implication: 'For all x, if x belongs to S–1, x has the property A'. This follows from the nature of *material implication*, i.e. conditional statement understood truth-functionally. Since we have assumed that the correlation between *S–1* and *A* holds, it follows that the universal implication is *confirmed* by the members of *S–1*, and that it is *not falsified* by the members of *S–2* or *S–3*. Now, if we take the (universal) material implication as an adequate formalization of natural laws, we are bound to eliminate the distinction between confirmation and non-falsification. (This follows from the fact that a material implication is true if its antecedent is false; and the members of *S–2* and *S–3* make the antecedent false.)[6] But then we are also bound to say, contrary to our initial assumption, that the law is valid not just for *S–1*, but also for *S*. — The only rational course of action is to *reject* the truth-functional interpretation of law-statements. This is in fact what has been done by those logicians who have developed a notion of 'strict implication', which does not give rise to above-mentioned 'paradoxes of confirmation'.

I have noticed that the point I am trying to make is not quite easy to grasp. Therefore I rephrase the above example in even simpler terms. Suppose that we have two laws, one (= *X*) that is valid for all pieces of metal and the other (= *Y*) that is valid for all pieces of iron (but not e.g. for pieces of silver or gold). Then it should be obvious that any conception claiming *Y* to be valid for *all* metals must be misleading.

The difference between statements like *X* and *Y* is clear in some domains (e.g. chemistry), but not in others (e.g. linguistics). In fact, what I am claiming here is that the notion of language universal, whether employed by Greenbergians or by Chomskyans, has so far suffered from an incapacity to distinguish between the two types of statement, the reason being the ambiguity surrounding the notion of material implication. For instance, let us accept the traditional tripartite classification into the isolating, agglutinative, and fusional (rather than 'inflectional') morphological types, and let us further assume that the agglutinative structure (but not the other two) is universally correlated with some syntactic characteristic *A*, resulting in the implication 'Aggl→A'. (Notice the similarity to the example with the three biological subspecies.) I submit that the majority of linguists would be willing to say — falsely — that this implication is valid for all languages (and thus represents a linguistic universal), because it is *confirmed* by agglutinative languages and it is *not falsified* by isolating or fusional languages. The reason for the present misunderstanding may thus be traced to the incapacity to distinguish between confirmation and non-falsification, an incapacity strongly promoted by the nature of material implication.

At this point I must add a proviso. Up to now, I have deliberately been speaking of *structural* laws (or 'laws of coexistence'), because it is to this category that implicational universals of the standard type obviously belong. It is a characteristic of such laws that they can be meaningfully said to be true only of those entities which make the antecedent true. The situation is different with respect to *experimental* laws (also called 'dispositions'), such as 'For all x, if x is a piece of metal, and if x is being heated, then x expands'. A linguistic counterpart might be e.g. 'For all x, if x is a natural language, and if normal children are exposed to x, then x will be learned'. It is obvious that the former law is true of all pieces of metal, not just of those that happen to be heated at some moment. Similarly, the latter law is true also of those natural languages to which no (normal) children are exposed. (We only need to generalize the obvious truth that the language spoken by a childless couple is a language.)[7] The difference between experimental laws and structural laws is as follows. Being heated (or being that which children are exposed to) is an external condition which may be equally well imposed upon any members of the homogeneous class constituted by the pieces of metal (or by the natural languages). By contrast, having a certain molecular structure (or a certain morphological structure) is one of those inherent properties which partition the non-homogeneous class of metals (or of natural languages) into subclasses.

I am arguing that, contrary to the prevailing opinion, implicational universals are not, as a general case, about *all* natural languages. Up to now, however, I have merely presented the outline of my argument. To make it fully convincing, I must first illustrate it in greater detail and then account for an important class of counterexamples.

For the sake of clarity, I shall make use of a fictive example. Let us consider the parameter of word-order with its six options VSO, VOS, SOV, SVO, OSV, OVS and the parameter of accessibility to relativization, i.e. subject > direct object > indirect object > oblique, with its four options S, S&DO, S&DO&IO, S&DO&IO&OBL; and let us assume, for the sake of argument, that if a language is VOS, then only the subject may be relativized, which yields the implication 'VOS→S'. This implication is made true by the constellations 'VOS&S', '~VOS&S', and '~VOS&~S'. The first of these admits only one option, which is, precisely, 'VOS&S'. Because '~VOS' equals all word orders *except* VOS, as well as the *lack* of any basic word order, the second constellation admits six options. And because '~S' equals all values on the accessibility-to-relativization parameter *except* S, as well as the *lack* of relativization altogether, the third constellation admits twenty-four options, thus:

$$\sim\text{VOS} \ \& \ \text{S} = \left\{ \begin{array}{c} \text{VSO} \\ \text{SOV} \\ \text{SVO} \\ \text{OSV} \\ \text{OVS} \\ - \end{array} \right\} \ \& \ \text{S}$$

$$\sim\text{VOS}\&\sim\text{S} = \left\{ \begin{array}{c} \text{VSO} \\ \text{SOV} \\ \text{SVO} \\ \text{OSV} \\ \text{OVS} \\ - \end{array} \right\} \ \& \ \left\{ \begin{array}{c} \text{S\&DO} \\ \text{S\&DO\&IO} \\ \text{S\&DO\&IO\&OBL} \\ - \end{array} \right\}$$

This information may be presented more conspicuously in the following tetrachoric table:

	VOS	~VOS
S	1	6
~S	–	24

The important thing is that we have here a law (or regularity) that is valid *only* for the VOS languages. For all we know, one SOV language, for instance, may be connected with S, another with S&DO, a third with S&DO&IO, a fourth with S&DO&IO&OBL, while a fifth may have no relativization at all. Thus, there is *no law* concerning the SOV languages (or the other non-VOS languages). Therefore, since our law is valid *only* for VOS languages, it would be obviously wrong to claim, with the majority of linguists, that it is valid for *all* languages (in spite of the fact that no language directly falsifies the law). It follows that implicational universals are in reality 'universals' only within quotes.

The same point may be made also in the following way. The implication 'VOS→S' expresses a law valid for the VOS languages, i.e. for what makes its antecedent true. This implication is logically equivalent to its so-called contraposition '~S→~VOS'. Interestingly enough, this implication does *not* express a law. The reason is that in this case, if the antecedent is true, the implication is made true by a 'heap' that implicitly contains 24 distinct options. Within such a heteregeneous conglomeration there can be no lawlike or nomic connections.

I submit that what I have been saying so far is relatively uncontroversial. But how is it possible, then, that the misunderstanding concerning the nature of implicational 'universals' has arisen, in the first place? Now we move to the second stage of my argument.

It is extremely important to realize that although implicational universals, as a general case, are not valid for all languages, a sizable number from among those implicational universals that have been proposed in the literature so far happen to be exceptions to this general rule. This explains, in my opinion, why the misunderstanding concerning the status of implicational universals has been so wide-spread.

Those implicational universals that happen to be genuinely universal in character contain predicates that, unlike VOS or S above, are *binary* in character. More precisely, 'binariness' means here the relation of opposition ('man vs. woman'), not just the relation of complement ('man vs. not-man'). It is a fact that up to now research has largely concentrated on binary predicates. It must be emphasized, however, that in spite of the popularity it enjoys, binariness is only a particular case.

Consider this statement: 'If a language has a nonzero morpheme for the singular, it has a nonzero morpheme for the plural'. The information expressed by this statement may be presented, and exemplified, by means of the following tetrachoric table (cf. Croft 1990: 68):

		A	~A
		singular morpheme	no singular morpheme
B	plural morpheme	Latvian	English
~B	no plural morpheme	———	Chinese

In this binary case '~A' and '~B', instead of being defined merely negatively (i.e. as what *A* or *B* is *not*), stand for something positive. Therefore in the present case, unlike in the previous one, both the implication 'A→B' and its contraposition '~B→~A' express a law. This means that, as can be seen from the tetrachoric table, both formulations of the law make a positive statement about *all* (types of) languages. For greater clarity, and also for the sake of comparison, the number of the options involved may be presented, as follows:

	A	~A
B	1	1
~B	_	1

It seems to me that such grammatical hierarchies as allow all (types of) languages to be classified informatively, i.e. not just based on the relation of complement, represent a generalization of the binary case. That they are universal, means therefore something more than the mere fact that they can be *formulated as* universal implications.[8]

Now we move to the third stage of my argument. If the law concerning the marking of grammatical number is formulated as 'If A, then B', it is a language like Latvian which makes the antecedent (as well as the consequent) true. If the law is formulated, by contraposition, as 'If ~B, then ~A', it is a language like Chinese which makes the antecedent (as well as the consequent) true. This brings out a curious fact. The law really wants to say that a language like English is the typical case: it has no morpheme in the ('unmarked') singular and has a morpheme in the ('marked') plural. But this 'typical case', which best exemplifies the law, makes the antecedent *false* in both formulations of the law. This is paradoxical, because entities which genuinely confirm a conditional statement should make its antecedent *true* (cf. above; also Johnson-Laird 1983: 54–63). Thus, 'If A, then B' (or 'If ~B, then ~A') is a somehow unnatural or inadequate way of expressing (markedness relationships exemplified by) the linguistic universal in question.

On closer inspection, this is not surprising. I have argued that this universal is genuinely valid for all languages, which means that it is an *unrestricted* (or non-implicational) universal in implicational guise. Therefore we must find a new and more natural way to express it. One way, mentioned already in Section 1.1, would be to view this universal as an exemplification of the following more abstract principle:

i) Primary needs of formal distinction are satisfied before secondary needs (= English).
ii) Therefore, if secondary needs are satisfied, primary needs are too (= Latvian).
iii) And, conversely, if primary needs are not satisfied, neither are secondary needs (= Chinese).

This formulation meets the desideratum of presenting the typical case as non-implicational and primary, and the less typical cases as (implicational and) entailed by the typical one. Notice that i) entails ii) and iii) just like 'A happens before B' entails both 'If B happens, A has happened' and 'If A has not happened, B has not happened'.

It may be added that there is an obvious difference between straightforward absolute universals (e.g. 'All languages have morphological structure') and what I have called absolute universals in implicational guise.

1.4. The Status of Linguistic Universals

I have argued above that the only genuine universals are *absolute* or *unrestricted* universals. (In fact, the very notion of 'language' presupposes the existence of such universals.) It follows that 'implicational universal' is likely to be a misnomer. Either it is about a subclass of all languages, which means that it is not universal. Or it is about all languages, which (normally) means that it is not implicational, except at the level of surface formulation. Only correlations between absolute or unrestricted universals would genuinely qualify as implicational universals (e.g. 'If, and only if, a language has vowels, it has consonants').

Now it might still be countered that implicational 'universals' are genuinely universal if taken in a *negative* sense, i.e. as 'constraints' on the notion of 'possible human language': they state that for *all* human languages something is *not* the case. This is true, but not very exciting. (Note 8 contains one such 'constraint'.) I must add that the logical justification for the 'search for constraints' has never become quite clear to me. Laws of nature are not viewed as 'constraints' (e.g. on the notion of 'possible metal'). Of course, they *can* be viewed in such a way, but what is the intellectual gain? And if (typical) laws of nature are not treated as constraints, why should laws of language be treated any differently?

Thus I repeat that the only genuine universals of language are (non-implicational) absolute universals. This may look like a discouraging conclusion. It is generally agreed that, in the present state of knowledge, there are rather few absolute universals of any theoretical interest. Therefore I ought to show next that my results do not undermine the very idea of language universals research.

If the sought-after universality cannot be found in particular correlations between linguistic properties, maybe it can be found in the common cause of these correlations. (I have already hinted at this possibility in what precedes.) Consider this statement: 'If a language has overt case-marking in direct objects, it has overt case-marking in indirect objects.' When looking at it, three things come to mind. First, this correlation is so similar to the one concerning overt case-marking in (in)transitive subjects, that the two must have a common explanation. Second, this explanation is in turn just a special case of the general principle governing differential needs of formal distinction. Thirdly, all statements about case-marking still apply only to a subclass of languages, i.e. they exclude the non-flectional languages. To include these as well, we have to raise the level of abstraction. It is certain (or can be 'predicted') that also in a language like Chinese, primary expressive needs must be satisfied before secondary ones. What we have to do is uncover the Chinese explananda, comparable to the correlations about case-marking, for this explanans.

Moreover, those working within the typological-functional framework have understood perfectly well that it is possible to provide 'deeper' explanations for (what they call) 'implicational universals'. Thus patterns of 'dominance' are explained in terms of the length (or 'heaviness') of grammatical elements while patterns of 'harmony' are explained in terms of analogy (or some underlying feeling of structural symmetry) (cf. Croft

1990: 53–63). I have the feeling that representatives of the typological-functional school tend to regard 'deeper' explanations of this kind as somewhat speculative. This probably stems from the in itself laudable wish not to claim more than can be warranted by the facts. (The difference vis-à-vis the Chomskyan approach is particularly evident here; cf. 2.4 below).

Personally, I think the interest should shift away from particular 'implicational universals' towards the larger problem of explaining universally valid facts about linguistic form and meaning. Such an approach has been outlined in Itkonen (1991b). It is true, of course, that in this field we may not be able yet to offer deterministic explanations. Therefore, if we had to give an exact formulation to our would-be explanations as they are right now, most of them would turn out to be of *statistical* form. Hammond *et al.* (1988) are probably not alone in treating statistical explanations with strong suspicion. This calls, again, for comments.

First, even regardless of the fact that the laws of particle physics, i.e. the 'basic laws of nature', are of statistical character, statistical explanations are considered as a legitimate type of explanation in current philosophy of science; and they have an obvious use in linguistics too (cf. Itkonen 1980 and 1983: 2.2.4, 6.1). Secondly, and more importantly, the statistical explanations of today may become deterministic explanations of tomorrow; with time, we may learn to fill in the gaps in our current explanations. (In fact, Einstein thought this could be done even with respect to the laws of particle physics.) Sometimes such a faith in future accomplishments is surely exaggerated, as in the case of historiography, where some die-hard determinists still insist that there *are* universally valid laws of human history, but we just do not *know* them (cf. Itkonen 1983: 95–96). Now, if this is allowed to happen in historiography, there is much more reason to let it happen in linguistics. Language is, after all, much better structured, and more easily surveyable, than the evolution of human societies on the globe.[9] — Thus I recommend the research program of explaining universal facts of language.

2. THE CHOMSKYAN APPROACH

Chomsky's version of the language universals research deserves close critical scrutiny. In the present context I shall be content to single out some of its major weaknesses. These should be added to those singled out in Itkonen 1991b.

2.1. The Types of Chomskyan Universals

Chomsky's universal grammar (henceforth to be abbreviated as 'PP') assumes the existence of innate parameters and principles. These are needed to explain the fact of language-acquisition, i.e. the 'fact' that children acquire their first language rapidly on the basis of degenerate and limited evidence (or even of 'no' evidence). The existential status of this 'fact' is extremely dubious. No precise meaning has ever been given to the claim that language-acquisition is rapid, rather than slow. The evidence that children encounter is neither degenerate nor limited. The claim that children know linguistic facts for which they have had *no* evidence rests on the assumption that children's intellectual capacities are extremely limited: they are quite incapable of perceiving any relationships (of similarity and difference) between the utterances they encounter; and they are almost equally incapable of retaining any memory of the (types of) utterances they have encountered.[10]

This assumption is contrary both to common sense and to traditional accounts of language-acquisition. It was assumed, e.g. by von der Gabelentz, Paul, de Saussure, Sapir, Jespersen, and Bloomfield, that children abstract certain patterns from the utterances they have heard and form new utterances on the analogy of these patterns (cf. Itkonen 1991a: 287–290, 299–304). The same view is implicit in Harris's (1951: 372) remark that "The work of analysis leads right up to the statements which enable anyone to synthesize or predict utterances in the language". It is interesting to note that in his dissertation Chomsky fully accepted this traditional account:

> A primary motivation for this study is the remarkable ability of any speaker of a language to produce utterances which are new both to him and to other speakers, but which are immediately recognizable as sentences of the language. We would like to reconstruct this ability within linguistic theory by developing a method of analysis that will enable us to abstract from a corpus of sentences a certain structural pattern, and to construct, from the old materials, new sentences conforming to this pattern, just as the speaker does (Chomsky 1975 [1955]: 131).

Nowadays PP contains such (presumably innate) *parameters* as '(syntactic) movement'. A language like English chooses the value '+' on this parameter, whereas a language like Japanese chooses the value '-', i.e. it has no (question) movement. The languages with the value '+' are in turn characterized by the *principle* of subjacency, which stipulates that movement may not cross more than one 'bounding' node. This principle, in turn, is the basis for the ('lower-level') *parameter* which says that, in addition to NP, languages may choose either S', i.e. COMP & S, or just S as a bounding node.

To give a few more examples, the 'head parameter' says that all languages are either 'head-first' or 'head-last', i.e. in NPs, VPs, APs, and PPs they have N, V, A, and P on the same side with the respect to the other material contained in the phrases (i.e. specifiers and complements). Moreover, any language must choose either '+' or '-' on the 'pro-drop parameter', i.e. it may or may not suppress the subject of a clause.

Subjacency is a principle which, at least on the face of it, is responsive to, and therefore falsifiable by evidence from different languages. By contrast, the 'projection principle', which stipulates that lexical structure must be represented at every syntactic level, is a theory-internal principle falsifiable, if at all, only in a very indirect way.

'Structure-dependency' is a general principle which says that linguistic operations are performed on (hierarchic) structures, rather than atomary units. It is an unrestricted universal. Subjacency, by contrast, is an implicational or restricted universal. It says '*If* a language has the value '+' on the movement parameter, then...'. A language like Japanese is taken to *confirm* this 'universal' because it does *not falsify* it (cf. 1.3. above).

2.2. PP and Explanation

As was noted above, PP is meant to explain the 'fact' of language-acquisition. Because of its innate character, however, the PP-type universal grammar itself is assumed to be unexplainable. It has often been pointed out that this is a kind of 'argument from laziness' (cf. Comrie 1981: 24, Hawkins 1985: 583): Before declaring something to be unexplainable, one should at least *try* to explain it.

Of course, Chomskyans have strongly rejected this interpretation, but their reasons for doing so remain confused. Hoekstra & Kooij (1988), for instance, refer to the 'theoretical foundation' that Chomskyans possess and their opponents presumably lack: whether or not a universal principle is decreed to be innate, results from 'theoretical argumentation'. Having made these unsubstantiated claims, Hoekstra and Kooij try to prove the correctness of their position more concretely, by showing that such phenomena as the Wh-movement can be given no functional explanation (pp. 45–52). But this just shows that they in fact *accept* Comrie's and Hawkins's argument: they do try to explain something, before declaring it to be innate. (It is a different matter that, quite obviously, they do not try hard enough.)

The conceptual confusion that continues to prevail in this area is strikingly illustrated by the following quotation:

> Before we can begin to evaluate explanations we have to know what it is that has to be explained. The position of generative grammar is, in respect, clear and consistent: what we have to *explain* are the principles underlying the child's ability to learn any language at all. A subset of these principles belongs to UG and is *innate* (Hoekstra & Kooij 1988: 49; emphasis added).

In reality, this presumably 'clear and consistent' position is unclear and inconsistent: Hoekstra and Kooij intend to explain precisely that which they, as opposed to Comrie and Hawkins, claim to be unexplainable, i.e. innate aspects of the language faculty.

If one wishes to apply the Davidsonian 'principle of charity' to what Chomskyans have been saying about innateness and (non-)explanation, they might be construed as saying the following thing: *if* the evidence for innateness is overwhelming, *then* the existence of (functional) explanations is so improbable that it is not worthwile to start looking for them. But of course, the evidence for innateness is far from overwhelming. For instance, the explanation of structure-dependency is self-evident. Linguistic structure reflects perceptual structure, in that they both exemplify the notion of what Jackendoff (1987: 249–251) calls 'headed hierarchy'. When I see a small boy eating a red apple, I see the smallness together with the boy and the redness together with the apple (rather than vice versa), and the NPs of my language (and, presumably, of any language) reflect this fact. Similarly, when I see a boy eating an apple, a man kissing a woman, and a dog chasing a cat, I see the boy together with the apple, the man together with the woman, and the dog together with the cat. The sentence-structures of my language reflect this fact: this is the only reason why I put the words *boy* and *apple* in the same sentence, instead of separating them by two sentences speaking about the man, the woman, the dog, and the cat. — The explanation given by Croft (1990: 179) in terms of 'iconic-distance hypothesis' is the same, except that he speaks of 'semantics', and not of 'perception' (as he should).

Personally, I do think that we have to do here with an 'argument from laziness'. As I noted in Itkonen (1991b), innateness and modularity merely serve as excuses for Chomsky to continue doing what he has always done, namely practise 'distributional analysis' on self-invented sample sentences which his intuitive knowledge of English deems as either correct or incorrect (cf. 2.6).

Sometimes it is said that even if a principle like subjacency cannot be explained, it explains something, namely why NPs can be moved in some ways, but not in others. A moment's reflection suffices to show that this is no (genuine) explanation. Let us assume, for the sake of argument, that the facts are as the subjacency principle claims them to be. Then several cases of ungrammaticality may be subsumed under this principle. However, this principle is a (mere) generalization out of, rather than an explanation of these cases.

An analogy will make this point clearer. Suppose that I have been given a large set of coloured figures, i.e. circles, rectangles, and triangles. I first notice that the first triangle is red and that the second triangle is red, and then I realize that all triangles are red. I have made a genuine generalization (= 'All triangles are red'), but it would not be appropriate to say that I have explained anything. In particular, I have not explained why this thing is red, if I have mentioned the fact that it is a triangle. A genuine explanation makes an at least implicit reference to causation (= why is it that all triangles, and not e.g. circles, have been painted red?). This is, very briefly, the reason why we do not speak of explanations in logic, although we do speak of generalizations and simplifications (cf. Itkonen 1978: 10.0). Of course, it is possible to 'psychologize' the subjacency principle and to claim that it is part of the machinery that makes us speak the way we do speak. But this is the 'virtus dormitiva' strategy. We can just as well 'explain' the fact of English plural-formation by saying that people form the plurals in the way they do, i.e. add the morpheme {S} with the three allomorphs /s/, /z/, and /ɪz/, because in their heads their have the mechanism which makes them form the plurals in the way they do, i.e. add the morpheme {S} with the three allomorphs /s/, /z/, and /ɪz/.

2.3. 'Universal Grammar of English Syntax'

"I have not hesitated to propose a general principle of linguistic structure on the basis of observations of a single language" (Chomsky 1980: 48). Those working outside the Chomskyan paradigm have found this type of statement rather preposterous. The medieval Modistae tried to construct a theory of universal grammar based on Latin, while the authors of the 17th-century Port-Royal grammar took French as the basis of their universal (or 'general') grammar (cf. Itkonen 1991a: 226–237, 261–269). It is generally agreed today that these two attempts were very largely failures. The failure did not consist in what the Modistae or the Port-Royal grammarians tried to do, but in how they did it: since they based their theory on observations of a single language, their data-base was just too narrow.

It looks self-evident that Chomsky is merely repeating the mistake of his predecessors. Surely it cannot be argued that the one-language approach to universal grammar is unjustified in one case (= Latin or French), but justified in the other (= English)? Amazingly, this is precisely what Chomsky's disciples have been willing to argue. This might be taken as a proof of Chomsky's infallibility within the paradigm that bears his name. That is, if his disciples had wished to build a plausible case for the one-language approach, they could have said, for instance, that Chomsky's statement should not be taken literally: although he occasionally claims to base his universalist hypotheses on observations of a single language, he is in reality making implicit use of his knowledge of other languages. Instead, the disciples have chosen to assert that when (and, apparently, only when) it is Chomsky who is using the one-language approach, it is fully justified. What they are really saying, is that Chomsky just cannot be wrong.

Hoekstra and Kooij (1988: 47) try to justify the one-language approach by arguing that the 'predictive power' of a universalist claim decreases as the set of languages constituting the data-base (i.e. the basis of prediction) increases. But this just shows that they have a confused notion of what science is about. Truth is a value in itself, predictive power is not. Suppose that I have to make a claim about all animals, and that I have restricted my data-base to mosquitos. (In zoology, this is not a realistic assumption, but as Hoekstra and Kooij are anxious to point out, in linguistics an analogous assumption is fully realistic.) Then I shall predict that all animals fly and have the size of approximately one inch. Of course, my claim has tremendous predictive power; but from the viewpoint of zoological theory, this fact does not, in itself, possess the significance that Hoekstra and Kooij attach to it.

Similarly, Cook (1988: 19) feels obligated to defend the one-language approach: "If the principle can be ascribed to the language faculty itself rather than to experience of learning a particular language, it can be claimed to be universal on evidence from one language alone." When you first read this sentence, it sounds plausible enough. But what it really says, is that if a claim is true, it does not matter how and why somebody came to assert it: 'If the principle can be ascribed to the language faculty itself, it can be claimed to be universal on evidence from fortune-cookies (or christal balls).' This may be so, but the only genuine question here is *how probable* it is that universalist claims based on one language, or on fortune-cookies, or on christal balls turn out to be *true*. And the answer is that in all three cases it is about equally *improbable*. Asserting this fact amounts to denying that there is a sharp dividing line between the 'context of discovery' and the 'context of justification'. This dichotomy was part of the philosophy of science in the 50s, but it has been abandoned since then. — In sum, it is preferable that claims about all languages should be based on as many languages as possible.

More recently, Chomskyans have been forced to abandon the one-language approach. An implicational or restricted 'universal' like the subjacency principle requires the knowledge of at least two languages (= English and Japanese). Otherwise subjacency would be falsely claimed to be an absolute or non-restricted universal. More generally, all parameters require the existence of at least two languages (with the values '+' and '-'.) These issues will be examined in the next subsection. Nevertheless, discussing the one-language approach was not wasted effort, because what it teaches about the Chomskyan approach remains true.

2.4. PP and Cross-Linguistic Evidence

With the 'principles-and-parameters' approach Chomsky's universal grammar seems to have opened itself to cross-linguistic evidence. Could this signal a rapprochement vis-à-vis the functional-typological school? Chomskyans promptly reject such a suggestion, and I think they are right to do so. This is due to the fact that, as I shall now proceed to show, cross-linguistic evidence stemming from the study of the world's languages plays a marginal role within PP.

For over twenty years, Chomsky's universal grammar contained no systematic treatment of case-systems. This was logical enough, because Chomsky was relying on the one-language approach, and his chosen language, i.e. English, has (practically) no cases. All this changed with the coming of PP. Chomsky realized that there are languages which differ from English in having a case-system. In a dramatic reversal of opinion, he now claimed that *all* languages have a case-system. Of course, languages like Chinese falsify this claim. Therefore the 'Case Theory', which is a 'module' of PP, assumes that the case-systems of all languages are *abstract* in the sense that they may or may not be 'morphologically realized'. Chinese just happens to be among the languages with a morphologically non-realized case system. — This is one more application of the 'depth vs. surface' distinction as it was practised in the 60s: The facts are complicated; thus, postulate a level where everything is simple, and call it 'depth'; call the facts 'surface', and forget about them.

Let us consider another example. Greenberg (1966 [1963]) noted certain less than perfect correlations (or 'tendencies') between the word orders within such pairs as determiner—noun, adjective—noun, noun—verb, and noun—adposition (i.e. pre- or postposition); and his followers have taken great pains to explain the lack of correlation, where this has seemed feasible. They need not have bothered, because Chomsky simplified everything with one stroke. The X-bar theory, or the phrase structure module of PP, contains a head parameter which flatly asserts that all languages exhibit perfect correlations between the word orders in NP, VP, AP, and PP: either they are 'head-first' or 'head-last' (cf. 2.1 here). What about those innumerable constructions in innumerable languages which do not obey this decree? Do they not falsify it? No, they are merely labelled (or branded) as 'marked' and set in opposition with the 'correct' constructions, which are called 'unmarked'. Thus markedness becomes, at the same time, an excuse for ignoring the cross-linguistic variation and a shield against falsification.[11] I cannot help feeling that Chomsky is here blaming languages for something for which he should blame himself. If he makes a claim which turns out to be falsified by a great number of languages, why punish these languages?[12] Notice also that while the X-bar theory (as part of PP) assumes Adjective Phrase to be a universal category, it is a well known fact that there are many languages which do not possess this category. Does this not worry the Chomskyans? No, they couldn't care less. — A similar criticism was voiced, maybe in slightly more diplomatic terms, by Comrie (1981: 7–8). To this day, it remains unanswered.

Let us consider one more example. According to Chomsky, each sentence begins with a complementizer, or COMP. (In more elaborate versions, the X-bar theory requires the pretheoretical notion of sentence to be construed as a 'COMP Phrase', with a mostly empty 'specifier', COMP as the 'head', and the sentence itself as a 'complement'.) The morphological realization of COMP in English may be *that* or *for*, but of course these words never occur in the beginning of a main clause. In fact, it is a nearly universal truth that main clauses never begin with a sentential particle. There is only one type of exception: some languages (including Finnish) employ a *question* particle. This is, then, the 'factual' (or 'cross-linguistic') basis for postulating the existence of a sentence-initial COMP. It seems quite obvious, however, that COMP, like any category employed by Chomsky, could have been postulated also without any evidence.[13]

I could go on, but I think the previous examples suffice to drive my point home. Cross-linguistic evidence plays a purely ornamental role within Chomsky's 'universal' grammar. (The need for putting 'universal' within quotes should have become evident by now.) First, most parameters require nothing beyond regimented knowledge of a couple of modern European languages. ('Pro-drop': Italian may suppress subjects, but English may not; 'Adjacency': French may put adverbs between verbs and objects, but English may not; 'Subjacency': Italian and French have S' as a bounding node, but English has S.) Second, insofar as parameters do refer to non-European languages, they still require no such knowledge as could not be acquired by spending one afternoon reading functional-typological literature. ('Movement': English vs. Japanese; 'Head parameter': taken, in a simplified form, from Greenberg [1966]). Third, cross-linguistic evidence is likely to be misleading anyway. (Chinese has no cases, but we must learn to ignore this fact and to see that it has Cases; Acehnese has no adjectives, but it still has Adjectives; etc.) The *de facto* ornamental nature of cross-linguistic evidence shows that the one-language approach (cf. 2.3) is still lurking in the background.

When Cook (1988: 17–20) claims that there is a difference between Greenbergian ('data-driven') universals and Chomskyan ('theory-driven') universals, she is right insofar as the former may and the latter may not be falsified by data. But she is quite wrong to argue that there is some sort of logical difference between the two. When she notes that a language in which a universal is not present does not disprove it, she is just reinventing the notion of (Greenbergian) implicational universal. Both frameworks contain unrestricted (or absolute) and restricted (or implicational) universals. As I argued in 1.3, only the former qualify as *genuine* universals.

2.5. What Would It Be Like to Learn Forms Without Meanings?

In this subsection I shall consider questions which are highlighted by the following quotation:

> Rationalists have typically construed primary data as syntactic in character. Chomsky, for example [sic], concedes that semantic information may facilitate syntax acquisition; however, he doubts that such information plays any role in determining how learning proceeds. Chomsky's reluctance to include semantic information, despite a number of studies that seem to indicate the relevance of such information, presumably stems from worries as to how the learner could possibly glean a sentence's meaning from the context of utterance (Matthews 1989: 61).

Chomsky inherited this formalist[14] attitude from the founders of North American structural (or 'taxonomic') linguistics. In his dissertation he rejected such 'mentalist' notions as 'ideas' and 'meanings', "for what were essentially Bloomfield's reasons", and claimed to be concerned, like Harris, merely with "the physical properties of utterances" (Chomsky 1975 [1955]: 86, 127, 63, n.1).

Bloomfield's hostility towards meaning was motivated by 'logical positivism', which was the prevailing philosophy of science in the 30s. It was required that "all scientifically meaningful statements...be translatable into physical terms — that is, into statements about movements which can be observed and described in coordinates of space and time" (Bloomfield 1936: 90); and it was not obvious to Bloomfield (nor is it to anyone else) how statements about sentence meanings could be so translated. Now, because the position of logical positivism on this issue is completely outdated today, it should be evident that Chomsky's reasons (which, to repeat, were originally 'Bloomfield's reasons') for concentrating on linguistic form alone are equally outdated.

During the heyday of logical positivism Carnap (1937) defended an analogous formalist program within the theory of logic. According to his 'principle of tolerance' (pp. 51–52), logic is nothing but a game played with meaningless formal units, with the consequence that everyone is free to invent his own rules of inference. This position, too, has been abandoned since then. It is interesting to note, however, that it had been anticipated, and refuted, by several philosophers of logic, notably Husserl (1913). He pointed out that there is a necessary connection between certain general categories of thought and the major expression-types of formal logic; and these, in turn, he regarded as being based on the major grammatical categories of natural language. Thus, the incorrectness of a sentence like *This tree is and* is not syntactic (or formal), but semantic in character: *and* is a sign of (or 'means') the operation of conjoining, but in this example nothing is conjoined to what precedes. The same applies to less drastic examples of 'syntactic' incorrectness as well. Husserl seems to be quite right (cf. Itkonen 1991a: 285–286). Closely similar views are being presented today e.g. by Halliday and Langacker.

As the quotation in the beginning of this subsection indicates, Chomskyans find the learning of meaningless forms unproblematical, and the learning of meaningful forms problematical. But they have reversed here the order of priorities. This issue deserves an extended discussion. In the present context I shall merely point out some of the most obvious flaws in the formalist, Chomskyan position.[15]

First, it is one of the best known results of experimental psycholinguistics that the learning of meaningless material is much more difficult than the learning of meaningful material. How can this fact be ignored in the context of language-acquisition?

Second, humans have an innate (sic) capacity to endow (results of) human actions with meanings. When children are said 'not to understand' something, their mind is not entirely blank (or concerned with pure form), but contains some vague or confused meanings. Similarly when adults first hear utterances of an unknown language, they attach to them some general meanings related to the speech situation, or at least to emotion and/or to sound symbolism. The same is true of hearing nonsense rhymes. The learning of *pure* form, if it ever occurs, is an abnormality.

Third, speaking is an action, consisting of several subactions. It is a conceptual truth that an action is made for a *reason*, which means that when someone does something, we can always ask *why* he did it. Thus when someone moves the verb or a Wh-word to the front of a sentence or suppresses a subject, there is always a reason for doing so (e.g., 'to make a question', or 'because it was not needed'). The Chomskyan framework requires us to envisage actions made for no reason at all.

Fourth, according to Chomsky's scenario, when the child hears a limited number of strings of sounds which we may identify as (physical) utterances of sentences of a certain language, he ('rapidly') learns this language. Oddly enough, it seems to have been generally overlooked that conditions that exactly meet these specifications obtain world-wide, *without* any language-acquisition taking place. I mean the exposure to non-native languages that children nowadays get when watching the TV or (preferably) listening to the radio. I know it for a fact that this exposure does *not* bring about language-acquisition, unless it is accompanied by some explicit teaching.[16] Thus mere sound (= 'pure form') is just not enough. What is required, in addition, is the (natural) context of use, i.e. precisely that aspect which Chomsky is anxious to suppress.

Fifth, it is generally agreed today that spoken languages and signed languages stem from a common faculty. The pervasive iconicity of the sign-languages (and in particular, of the pointing signs) makes it impossible even to entertain the idea that those who are learning a sign-language would be learning 'pure form'. But then, because of the common ancestry, those learning a spoken language cannot be learning 'pure form' either.

Sixth, I finally turn to the Matthews-quotation. The first thing to notice are the curious difficulties that Chomsky experiences in trying to figure out how the child manages to 'glean meaning from the context of utterance'. The associationist learning theory already provided the adequate answer: When a child sees a dog and hears *dog*, he has learned that *dog* means 'dog'. If you ask how this is possible, I answer that children are just made that way, i.e. they are *innately equipped* to make associations of this kind; and I do not mean this as a joke. Surely the representatives of associationism (including Aristotle and Hume) have always claimed that there is an innate basis for making associations. The meanings of verbs like *chase* are learned similarly, i.e. by associating the verb-forms with something (here: actions) occurring in the context of utterance. Learning the sentence-meanings is just a matter of learning to associate e.g. ontological 'thing—action—thing' triples with linguistic 'noun—verb—noun' triples (e.g. *dogs chase cats*, meaning 'dogs chase cats').

More importantly, however, the Matthews-quotation reveals the underlying rationale of Chomsky's entire formalist program. Chomsky admits, although reluctantly, the *de facto* importance of semantics, but he dismisses it, because he does not know how to handle it. Notice what this really means. There are two positions here: P–1 = 'Language-acquisition requires syntax *and* semantics' and P–2 = 'Language-acquisition requires *only* syntax'. Chomsky admits that P–1 is true. P–1 excludes P–2, which means that Chomsky must admit that P–2 is false. However, he does not know how to handle (i.e. how to formalize) P–1. Therefore, he rejects P–1 (which, to repeat, he knows to be true) and chooses P–2 (which he knows to be false).[17]

2.6. PP Is Not a Falsifiable Theory

It is one of the putative virtues of PP (as well as of its predecessors) that it makes very *specific* claims: because of their specificity, it can be immediately seen whether the claims have been falsified or not. Moreover, rival approaches are taken to be inferior to PP precisely because the claims they make are less specific. We have already seen that this argument has no force. In this subsection I shall explain in somewhat more detail why this is so.

In the 60s Chomsky made use of two principal 'immunization strategies' to avoid falsification: First, if a fact could not be accommodated, it was relegated to *performance*. Second, if the refractory fact belonged to competence too obviously for the first strategy to apply, it was explained away as a *surface* aberration; the depth still conformed to the theory. These strategies are still with us, but their field of application has considerably expanded.

Nowadays it is mainly the notion of *markedness* which is resorted to, in order to ward off the spectre of falsification. This happens in two steps. First, it is stipulated that linguistic phenomena divide into two different types, namely *core* and *periphery*, and that the PP-type universal grammar concerns itself only with the former. Thus each language is taken to contain a 'core language' which conforms to PP. Facts belonging to the 'periphery' cannot falsify PP: "The more something departs from UG the more it is *marked*" (Cook 1988: 81; emphasis added).[18] Insofar as Greenberg-type universals do not agree with PP, they are declared to be part of the periphery. It seems clear that the 'core vs. periphery' distinction is a generalization of the older 'competence vs. performance' distinction.

Second, there is markedness also within the core. Just as formerly in connection with the competence, now also in connection with the core there are certain facts which, although contrary to the theory, can only with difficulty be thrown into the garbage can (whether you call it 'performance' or 'periphery'). This is where the notion of 'parametrized variation', combined with markedness, comes in. PP wants to make unrestricted claims about all languages: for instance, all languages have the Wh-movement, or all languages have adjectives. It turns out, however, that such claims are false. They are made to appear true, however, by assuming that there are 'parameters' on which the languages conforming to the original claims of PP have the value 'unmarked' while the other, offending languages have the value 'marked'.

The immunization strategies which I have spoken about here presuppose another such strategy, namely restricting the data to conscious *intuitive judgments* about grammaticality. This is often denied, and it is claimed, instead, that PP may be either confirmed or falsified by many different kinds of data. We will see now, however, that such claims are unfounded.

Let us consider data from language-acquisition. Slobin (1973) suggested that cross-linguistic comparisons of the relative difficulty with which children learn different types of construction might reveal what is universal and what is less so in the language faculty. For some time, psycholinguists working within the Chomskyan paradigm took up this idea. It turned out, however, that children do not acquire their first language in the way predicted by Chomsky. Does this mean that Chomsky's theory was falsified? Of course not. Goodluck (1986) hit upon the lucky idea that children just have *wild grammars*, i.e. grammars disagreeing with PP; and since then it has become customary to warn that data from language-acquisition is 'potentially misleading' and should be treated with extreme caution. When put in plain language, Chomskyan child psychologists are sending us the following message: If Chomsky is wrong, blame it on children.

This reaction was only to be expected, because some twenty years earlier Chomsky had already ruled out the use of experimental psycholinguistic evidence. That is, in the mid-sixties it looked for a moment as if psycholinguistic experiments had established the psychological reality of deep structures and transformations. Chomsky enthusiastically accepted this confirmatory evidence.[19] But when subsequent experiments invalidated the assumption of deep structures and transformations, Chomsky coolly discarded this disconfirmatory evidence. In so doing, he committed the cardinal sin of any scientist: to accept the evidence only so long as it is profitable to do so. When Chomsky (1986: 36–37) claims that "evidence... could come from many different sources [including] perceptual experiments, the study of acquisition...or language change", we can be sure that he does not really mean what he says.

"Criteria of refutation have to be laid down beforehand: it must be agreed which observable situations, if actually observed, mean that the theory is refuted" (Popper 1963: 38).[20] We have seen that Chomsky's theory has never adhered to this fundamental methodological requirement. Now, as against this, it might be — and in fact has been — argued that when TG (with its own successive instantiations) was replaced by GB, and GB by PP, this means that the earlier theory was *falsified* by the later one. Does this not show, then, that Chomsky's theory is falsifiable? No, it does not. It shows at most that *Chomsky has been free to change his mind.* That is, there is no systematic justification for the changes involved in the development 'TG > GB > PP'. Rather, the changes have been exclusively motivated by Chomsky's personal tastes. No one else has ever falsified the smallest bit of PP or of any of its predecessors. Only such 'falsifications' as have been sanctioned by Chomsky have influenced, and will influence, the development of that linguistic theory which is currently being instantiated by PP. (The text books on Chomskyan linguistics studiously avoid mentioning this vital piece of information.) Science is an intersubjective and cooperative undertaking. Measured by this criterion, PP hardly qualifies as a scientific theory.

NOTES

1. Comments made by Catherine and Jon Ringen as well as by Maria Vilkuna have prompted me to express myself more clearly in a number of passages.

2. That the two sentences are equally possible does not mean, of course, that they are true with equal probability. One type of cause has (typically) only one effect, but one type of effect may have an indefinite number of causes. To see the point, compare these sentences: 'If anyone has been dropped without a parachute from the president's airplane by a group of half-naked eskimos, then he is dead' vs. 'If anyone is dead, then he has been dropped...'.

3. However, the effect may be the cause of, and thus explain, the process of hypothesizing about its cause: 'Why do you assume that it has rained?' — 'Because the streets are wet.'

4. This is not meant to be a water-tight truth. Prima facie at least, it is difficult to interpret e.g. some word-order universals in this light. Cf. 'If a language has the VSO order, it has SVO as an alternative'.

5. Croft (1990: 53) assumes, on the contrary, that *B* is dependent on *A*. Yet he also assumes (p. 54), together with Greenberg, that *B* is 'dominant' with respect to *A*.

6. In fact, anything that does not belong to *S-1*, pieces of paper as well as galaxies, makes the antecedent false and the implication true.

7. My use of illustrative examples is unsystematic insofar as in one context I compare languages to metals qua natural kinds and in another to pieces of metal. I do not think, however, that this variation is of any consequence.

8. Notice that even a sentence which is (meant to be) valid for a single language may be formulated as a universal implication; cf. 'For all languages x, if x has a word *mies* meaning 'man', x has a word *nainen*, meaning 'woman',' which is true of Finnish only.

9. I still subscribe to the view, expressed in Itkonen (1982), that if there are any deterministic laws of linguistic *change*, their level of abstraction is much higher than that of standard 'diachronic explanations'.

10. Chomsky regards 'utterances' as *formal* entities. Thus, on this view, the input that children get consists of meaningless strings of sounds which are taken out of context. The inadequacy of this conception is evident from the fact that, in the beginning, children learn the meanings of *only* those words whose referents occur in the speech situation. More generally, it is self-evident that understanding the meaning of sentences is a prerequisite (and the motive) for learning their form. By making the opposite assumption, Chomsky manages to make language-acquisition appear as an impossible

task, i.e. a task that can be accomplished only with the aid of intricate innate machinery; cf. 2.5 below.

11. We should not think that the notion of markedness is discredited by the fact that it may be misused in this way. It may, however, be questioned on other grounds. In particular, it is a mistake to think that markedness is an explanatory concept. It is, rather, a cluster concept, i.e. it combines form, distribution, and frequency. These three criteria do not always coincide; and when they do, this is a fact which needs explanation.

12. No natural scientist could afford to take an equally cavalier attitude vis-à-vis the facts. It would be beside the point to argue that in the natural sciences too a paradigmatic theory may have to accept the existence of 'anomalies'. What we have here are not a few anomalous or recalcitrant facts, but massive, across-the-board falsification.

13. Cross-linguistic evidence almost comes to look as an unnecessary complication. Everything would be much simpler if the 'surface' were *totally* wrong, and *totally* divorced from the 'depth'.

14. In the present context *formalist* means 'concerned with form, as opposed to meaning', and not 'making use of formal methods'.

15. It may be good to add that operating with 'theta-roles' does nothing to mitigate Chomsky's formalist position because these 'roles' merely replicate syntactic relations. Thus, in a sentence like *John suffered an injury* *John* is 'agent' and *injury* is 'patient'; cf. Ravin 1990: chap. 3.

16. Ever since he was born, my son has heard each day about 100 English utterances, either spoken or sung. After a couple of years, this data could hardly be called 'limited'. Yet he started to learn English only at the age of seven, after some teaching.

17. It is not just a pun to say that form lends itself quite naturally to formalization, i.e. much more naturally than use-in-context, which may contain various kinds of semantic cues. Sticking to what is most easily formalizable, come what may, is another instance of the 'argument from laziness'.

18. Try to imagine a comparable principle in physics: 'The more something departs from our current theory of sound-waves the more it is (e.g) weird'.

19. As Gaberell Drachman told me in Salzburg, July 1977, "Chomsky came to Chicago waving the flag. 'They have proved it, they have proved it', he exulted."

20. The word *observable* may have to be replaced by *intuitional* in some contexts.

REFERENCES

Bloomfield, Leonard. 1936. 'Language or ideas?'. *Language*.

Carnap, Rudolf. 1937. *The logical syntax of language*. London: Routledge.

Chomsky, Noam. 1975 [1955]. *The logical basis of linguistic theory*. New York: Plenum Press.

_____. 1980. 'On cognitive structures and their development'. M. Piattelli-Palmarini (ed.): *Language and learning*. London: Routledge.

_____. 1986. *Knowledge of language*. New York: Praeger.

Comrie, Bernard. 1981. *Language universals and linguistic typology*. London: Blackwell.

Cook, V.J. 1988. *Chomsky's universal grammar*. London: Blackwell.

Croft, William. 1990. *Typology and universals*. Cambridge UP.

Goodluck, Helen. 1986. 'Language acquisition and linguistic theory'. P. Fletcher & M. Garman (eds.): *Language acquisition*. Cambridge UP.

Greenberg, Joseph. 1966 [1963]. 'Some universals of grammar with particular reference to the order of meaningful elements'. J. Greenberg (ed.): *Universals of language*, 2nd ed. Cambridge MA: The MIT Press.

_____. 1978a. 'Typology and cross-linguistic generalization'. J. Greenberg (ed.) *Universals of human language*, vol. I. Stanford UP.

_____. 1978b. 'Diachrony, synchrony, and language universals', *ibid*.

Hammond, Michael, Edith Moravcsik & Jessica Wirth. 1988. 'Language typology and linguistic explanation'. M. Hammond et al. (eds.): *Studies in linguistic typology*. Amsterdam/Philadelphia: Benjamins.

Harris, Zellig. 1951. *Methods in structural linguistics*. Chicago UP.

Hawkins, John. 1985. 'Complementary approaches in universal grammar'. *Language*.

Hoekstra, Teun & Jan Kooij. 1988. 'The innateness hypothesis'. J. Hawkins (ed.): *Explaining language universals*. London: Blackwell.

Husserl, Edmund. 1913. *Logische Untersuchungen*. 2nd ed. Tübingen: Niemeyer.

Itkonen, Esa. 1978. *Grammatical theory and metascience*. Amsterdam: Benjamins.

_____. 1980. 'Qualitative vs. quantitative analysis in linguistics'. T. Perry (ed.): *Evidence and argumentation in linguistics*. Berlin: de Gruyter.

_____. 1982. 'Short-term vs. long-term teleology in linguistic change'. P. Maher et al. (eds.): *Papers from the 3rd international conference on historical linguistics*. Amsterdam: Benjamins.

_____. 1983. *Causality in linguistic theory*. London: Croom Helm.

_____. 1991a. *Universal history of linguistics: India, China, Arabia, Europe*. Amsterdam/Philadelphia: Benjamins.

_____. 1991b. 'Two notions of universal grammar'. *SKY: The Yearbook of the Linguistic Association of Finland.*

Jackendoff, Ray. 1987. *Consciousness and the computational mind.* Cambridge MA: The MIT Press.

Johnson-Laird, P.N. 1983. *Mental models.* Cambridge UP.

Matthews, Robert. 1989. 'The plausibility of rationalism'. R. Matthews & W. Demopoulos (eds.): *Learnability and linguistic theory*. Dordrecht: Kluwer.

Moravcsik, Edith. 1978. 'Introduction'. J. Greenberg (ed.): *Universals of human language*, vol. 4. Stanford UP.

Popper, Karl. 1963. *Conjectures and refutations*. London: Routledge.

Ravin, Yael. 1990. *Lexical semantics without thematic roles*. Oxford UP.

Slobin, Dan. 1973. 'Cognitive prerequisites for the development of grammar'. Ferguson, C.A. & D. Slobin (eds.): *Studies of child language development*. New York: Holt, Rinehart, and Winston.

Stegmüller, Wolfgang. 1974. *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*, vol. I, 2 edn, Berlin: Springer.

Wason, P.C. & P.N. Johnson-Laird. 1972. *Psychology of Reasoning*. Cambridge: Harvard UP.

Address: University of Turku
Department of General Linguistics
Henrikinkatu 4
SF–20500 Turku 50, Finland