# Concerning the generative paradigm

Esa Itkonen

Abstract

This paper is a critical evaluation of the generative paradigm. The following aspects of this paradigm will be singled out for discussion: The definition of language, the concept of language acquisition, the concept of universal grammar, and the methodological notions of falsification and explanation

## 1. Preliminary remarks

According to a wide-spread opinion, generative linguistics is the dominant linguistic paradigm, at least in the United States. Typological-functional linguistics and cognitive linguistics might qualify as rival but less dominant paradigms. There are several publications that extol the virtues of the generative paradigm, most recently Pinker (1994), which simply identifies 'linguistics' with 'generative linguistics'. By contrast, there are nowadays relatively few publications which subject generative linguistics to explicit critical scrutiny. The reason for this state of affairs is not that generative linguistics contains no aspects open to criticism. Rather, representatives of other schools seem anxious to maintain what might, depending on one's point of view, be called either 'peaceful coexistence' or 'balance of terror'.

In the spring 1994 there was in the e-mail box 'Linguist List' a lengthy discussion concerning 'mainstream linguistics'. This discussion was started by the observation that people who had to criticize the generative paradigm preferred to stay anonymous. This was interpreted as reflecting the opinion that open criticism of 'mainstream linguistics' might jeopardize one's career prospects and possibilities for publication. Although such a view was heatedly denied by representatives of the generative paradigm, at least my own experience confirms it. In several informal meetings I have found out that linguists agree with some or even all aspects of the criticism that will be presented below; they just do not want to say it publicly. In my opinion, this is reason enough to present this criticism.

Before doing so, however, one qualification is in order. It goes without saying that Chomsky is a brilliant syntactician. If his syntax had a more straightforward connection with linguistic data, it would be a monumental achievement. But even if this were the case, the wider theoretical framework within which he wishes to embed his syntax would remain open to serious criticism.

## 2. Language

### 2.1. The definition of 'language'

Within the generative paradigm it was for a long time held to be a "truism that language is an infinite set of sentences" (Chomsky 1969: 57). Far from being a truism, however, the decision to view language in this way had far-reaching consequences concerning the overall conception of linguistics. Defining language as a set made it natural to think of sentences as objects. From this it was only a small step to viewing sentences as physical objects. And it goes without saying that physical objects have to be investigated by methods of natural science.

This reificatory tendency was reinforced by obliterating the distinction between natural languages and artificial languages whose 'sentences' are just strings of a's and b's. In contradistinction to natural languages (as well as to the languages of logic and mathematics), such ab-languages cannot be used in any meaningful sense of the word. Consequently there can be no rules (of use) connected with these

languages, and they can only be defined as (infinite) sets of sentences, generated or not by grammars each of which in turn consists of a (finite) set of grammatical rules. (Notice the difference between 'rules of language' and 'rules of grammar'.)

Artificial languages have exerted another type of analogical influence as well. They were created in order to get rid of the 'messiness' of natural language. For instance, while natural-language sentences may be more or less grammatical (with clear cases of both grammaticalness and ungrammaticalness), formal logic assumes that all formulae are either well-formed or non-well-formed; and two-valued formal logic assumes that all well-formed formulae are either valid or invalid. In a curious 'backward motion' Chomsky's syntax applies to natural language this two-valued discreteness which was achieved by 'going away' from natural language in the first place.

Now, there is no obvious reason to define natural languages as sets of (grammatically discrete) sentences because, unlike the ab-languages, they are quite incontestably spoken (and written) in accordance with socially valid rules. Occasionally this was realized also within the generative paradigm: "A natural language is properly viewed in the good old way: viz., as a system of conventions for the expression of communicative intentions" (Fodor 1975: 106). Such flashes of common sense remained without any permanent influence, however. (Notice that in the present context the terms 'rule', 'convention', and 'norm' are interchangeable.)

By the mid-80's Chomsky's definition of language underwent a change. Now there are assumed to be two types of language, namely externalized language (= 'E-language') and internalized language (= 'I-language', formerly 'competence'), of which the former is a mere artifact while the latter constitutes the genuine subject matter of linguistics (cf. Chomsky 1986: 15-50). This terminological shift creates the illusion that I-language has somehow become transparent enough to be investigated directly. The fact has now been forgotten that I-language is a highly hypothetical entity, conceived differently by every school of linguistics. As a matter of fact, the primary access to (unconscious) I-language can only be provided by conscious intuitions about sentences that belong to a public or social language (roughly, E-language). By thus obliterating the distinction between conscious and unconscious, Chomsky gives a misleading picture of how linguistic investigation actually proceeds, with harmful consequences.

This can be shown in more detail as follows. I-language is claimed to be a procedure that generates structural descriptions, also called 'expressions of the language' (Chomsky 1992:1). It might seem possible to interpret these entities either as descriptions of or simply as identical with the class of grammatical sentences of L. However, this possibility is explicitly ruled out: "The class so defined has no significance. The concepts 'well-formed' and 'grammatical' remain without characterization or known empirical justification" (ibidem, pp. 63, n. 7).[1] This statement has interesting implications. For instance, Chomsky (1992: 59) analyzes these two sentences:

John wondered what stories about us we had heard
*John wondered what stories about us we had told

The star is presumably meant to indicate that the latter sentence is ungrammatical. But notice that according to Chomsky himself, he has no 'empirical justification' for making this judgment. (And since the social E-language is a mere artifact, he has no reason to assume that anyone else will share his - unjustified - judgment.) The primary data of linguistic descriptions have now been located in some sort of undefinable limbo. This is graphically shown by the fact that most of Chomsky's (1992) 'examples' represent some intermediate level between sentences and grammatical descriptions, i.e. they are 'sentences' containing brackets and other descriptive devices.

2.2. The shortcomings of 'I-language'

It seems likely that Chomsky has decided to abandon the notion of grammaticalness because he has finally realized that it is an inescapably normative notion; and the normativity of linguistic data is incompatible with the assumption that linguistics is a natural science (because e.g. physical data are inherently non-normative) (cf. Itkonen 1978: 122-131, 175-190; 1983a: 54-61, 75-76). We have seen that this forces Chomsky to adopt a self-inconsistent position (= he rejects the notion of grammaticalness on which he relies). The same can be said, more generally, of his conception of I-language. At least the following contradictions may be pointed out.

On the one hand, I-languages are "real elements of particular minds/brains, aspects of the physical world" (Chomsky 1986: 26). On the other, the rules which constitute an I-language are regarded as comparable to the rules of chess (p. 27). This does not make sense, however, because the game of chess is constituted by one set of rules which is common to all chess-players. It would be almost comical to assume that there are millions of games of chess (as "elements of particular minds/brains"). It is an entirely different matter that every chess-player must have internalized the rules of chess; and such internalizations may well vary among themselves. (This distinction between social-normative and individual-psychological is the cornerstone of Itkonen 1983a).

On the one hand, the rules of I-language are not normative: "they entail nothing about what [one] ought to do" (p. 241). On the other, as we just saw, they are comparable to the rules of chess. But this is a contradiction because it is not reasonable to deny that the rules of chess determine how one ought to play. Thus, in their written form the rules of chess do not just describe behavioral regularities.

On the one hand "rules [i.e. rule-sentences] are not descriptions of behavior or of regularities in behavior", because everyone can decide to break them if he so wishes (p. 231). On the other, rules are "appropriate to describe...the way a spider spins a web" (p. 239). Again, this is a contradiction, because it is not reasonable to assume that a spider can decide to break the 'rules' of its behavior.

Thus, the notion of I-language turns out to be hopelessly confused. Chomsky would have avoided these problems if he had stuck to his original conception of language. As he rightly notes (p. 19), Bloomfield defined a language as (one version of) an E-language, i.e. "the totality of utterances that can be made in a speech community". This was also Chomsky's position in his dissertation: "The grammar thus gives a theory of these utterances...we have, at any time, only a finite corpus of utterances out of an infinite set of [possible] utterances" (Chomsky 1975a [1955]: 77-78). Consonant with this attitude, Chomsky at that time outlined a theory that "will rule out mentalism for what were essentially Bloomfield's reasons, i.e. its obscurity and inherent untestability" (p. 86). Rejecting mentalism of course entails rejecting any notion of I-language. Therefore it is quite interesting to see that Chomsky nowadays claims I-language to have been the topic of his dissertation ("although the term was not used"; Chomsky 1986: 48-49, n. 17). In conformity with Orwellian Newspeak, history is now being rewritten in such way that rejection of X has become acceptance of X. Within the generative paradigm, this might be called 'Orwell's problem'.

### 2.3. Private vs. public language

Wittgenstein's private-language argument may be used to show that Chomsky's language-conception is untenable (cf. Itkonen 1978: 4.0; Itkonen 1983a: 5.1.2-3, 5.1.5). Kripke (1982) too has presented this type of criticism, and Chomsky (1986: 5.1) has responded to it. In the discussion between Kripke and Chomsky there are two points that have not received the attention they deserve. First, when Wittgenstein speaks of 'private rules', he has in mind rules that are invented and followed consciously. This means that in this context it is a mistake to speak of the rules (or rather, 'rules') of I-language. Fodor (1975) too has failed to see this (cf. Itkonen 1983a: 5.1.4).

Second, the entire discussion centers on Wittgenstein's claim that a person putatively speaking a private language cannot know whether or not he has made a mistake; and because the notions of

language and rule presuppose the possibility of making a mistake, there can be no private language. Let us make this a little clearer. Let us assume that at this very moment I am going to (consciously) use a word that I myself have privately invented. My use of the word, i.e. what I mean (or intend to mean) by it, is based on my particular memory of how I have used it in the past. Maybe I wish to check this particular memory to make sure that I am not mistaken. But the only check I can rely on is this same memory; and of course it is no independent or genuine check; in fact it is no check (or basis for testing) at all. Therefore any 'private' rule-application that seems correct to me will be correct, which means that the notion of a private rule-application, and thus of a private language, collapses.[2] Genuine checks are provided only by other people's memories, and more generally by their intuitions about the correct use of language. Of course, there is no guarantee that these are always trustworthy. But at least they offer the possibility of genuine testing; and possible testing is certainly preferable to impossible testing (represented by my exclusive recourse to my own memories or intuitions).[3] This is nothing but the requirement of 'independent evidence', which is the cornerstone of scientific thinking. I find it odd that its connection with the private-language argument has generally not been understood.

This argument applies to any language that contains rules that can be brought onto the level of consciousness. All natural languages are of this type, given that their word-meanings, for instance, are based on potentially conscious rules. German contains rules to the effect that Berg and aber mean 'mountain' and 'but', respectively, while the rules of Finnish determine that these meanings are expressed by vuori and mutta, respectively. If, to express the meaning 'mountain', a German speaker uses such word-forms as aber, vuori, mutta, or mountain, he breaks a rule of the German language. Notice, however, that one can also decide to break purely formal rules of language (as Chomsky's own example Furiously sleep ideas green colourless demonstrated long ago).

It should also be clearly understood that the private-language argument applies only secondarily to language. First and foremost, it is meant to apply to knowledge in general. The real target of Wittgenstein's critique is the Cartesian idea that knowledge is primarily subjective. Wittgenstein argues that knowledge is primarily intersubjective, regardless of whether it is knowledge about observable events or knowledge about rules (or norms) accessible to intuition. Why? - because in the domain of genuinely subjective knowledge, expressed by my private language, I cannot know whether or not I am mistaken.[4]

The private-language argument shows that the rules of language are necessarily public, which means that language is primarily social (or intersubjective). Chomsky (1975b: 71) denies this: "As for the fact that the rules of language are 'public rules', this is indeed a contingent fact" [emphasis added]. It is interesting to note that Chomsky's non-social view of language is facing problems even within the generative paradigm. Recent studies have shown that infants display social behavior very early (cf. Butterworth & Grover 1988), and this could be taken to mean that at least some social concepts are innate. Jackendoff (1992: Chapter 4) embraces this conclusion whole-heartedly, and assumes the innateness of such social concepts as 'person', 'request', 'transaction', and 'ownership'. The interesting thing is that Jackendoff, as a "deeply committed Chomskian" (p. 53), cannot avoid coming into conflict with the central philosophical thesis of the generative paradigm. In accordance with his non-social or biological conception of language, Chomsky (1976: 183) claims that the linguist studies language exactly as the physicist studies physics, namely "taking humans to be 'natural objects'". But now the following contradiction emerges: On the one hand, the innate concepts of person and request are social. On the other, the innate linguistic expressions for these concepts (i.e. grammatical person and imperative) are non-social.

In other words, we are required to assume the existence of forms for e.g. request (or question) which have no connection with their use as request (or question). This makes no sense because 'X is a form for Y' means the same thing as 'X may be used to do Y'. (The connection with the private-language argument resides in the fact that requests and questions, being directed to other people, necessarily

point beyond the 'private' sphere of the one who 'owns' the forms for request/question.)[5] Notice another interesting implication of the Jackendoff-type innateness: When you meet a person, he is a social being; but as soon as he starts to speak, he miraculously changes into a natural, non-social being.

It is almost redundant to add that Chomsky's view of linguists taking speakers to be natural objects is directly contradicted by his own descriptive practice. The only speaker Chomsky has ever investigated is himself; that is, all he has ever done qua descriptive linguist is (self-reflectively) investigate his own conscious intuitions about various English sentences. Natural objects lack consciousness; a fortiori, they lack the powers of self-reflection. Therefore it is wrong to say that Chomsky qua linguist has investigated himself qua speaker as a natural object (for discussion, cf. Itkonen 1978: 81-86, 117-121; 1983a 227-248).

2.4. Form vs. meaning-and-form

For Chomsky, language is not only a non-social, but also a formal entity. This point may be clarified by the following quotation:

"Rationalists have typically construed primary data as syntactic in character. Chomsky, for example, 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 attitude from the founders of North American structural (or 'taxonomic') linguistics. In the same context in which he rejected mentalism, "for what were essentially Bloomfield's reasons", he also rejected such notions as 'ideas' and 'meanings' (Chomsky 1975a [1955]: 86); and he proposed to practise Harris-type distributional analysis, concentrating on "the physical properties of utterances" (p. 63, n.1, p. 127). The view of sentences as 'physical events' has remained the same ever since (see e.g. Chomsky 1986: 26).[6]

Bloomfield's hostility towards meaning was motivated by 'logical positivism', which was the prevailing philosophy of science in the 30's. It was required at the time 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 The 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 1991: 285-286). Closely similar views are held today by representatives of cognitive linguistics.

The Matthews-quotation above shows that Chomskyans find the learning of meaningless forms unproblematical, and the learning of meaningful forms problematical. Here they have reversed the order of priorities. This issue will be discussed more extensively in Subsection 3.3.

## 3. Acquisition of language

### 3.1. 'Poverty of the stimulus'

Chomsky's universal grammar is invoked to explain the 'fact of language-acquisition'. That is, language is claimed to be acquired 'very rapidly' on the basis of 'limited' and 'degenerate' data. Presumably the 'stimulus' which the child comes across is so 'poor' that he could not learn his native language so rapidly (or could not learn it at all), if he were not aided by a specifically linguistic innate faculty. Thus, the need for an innate universal grammar is justified by this 'poverty-of-the-stimulus' argument.

Curiously enough, it has proved quite difficult to reach any consensus on the existence of the very fact that the Chomsky-type universal grammar is summoned to explain. In other words, it is denied that language is acquired rapidly (cf. Sampson 1980: 114), and that the data encountered by the language-learning child is either degenerate (cf. Labov 1972: 203) or limited (cf. Schlesinger 1975: 219). In fact, it is quite evident that, contrary to what Chomsky assumes, in a normal language-learning situation the data encountered by the child cannot under any reasonable interpretation be considered as limited.

It has turned out, however, that there are some situations, originally not envisaged by Chomsky, in which the input given to the child is indeed limited in the sense of not being representative of a full-fledged language. For instance, a deaf child of hearing parents creates on his own a gestural 'home-sign system' that goes beyond the input that he receives. However, contrary to what is the case in a typical sign language, all signs of a home-sign system are iconic (cf. Goldin-Meadow & Mylander 1990), which shows that there still remains a connection with prelinguistic behavior. Similarly, children who create a creole on the basis of a pidgin go beyond the input they receive. Creoles exemplify to an extraordinarily high degree the principle 'one meaning - one form' (cf. Bickerton 1981: 145-146, 189-190). This principle is certainly innate, but there is no reason to assume that it is exclusively linguistic (cf. Markman 1991: 100-101).

Since the beginning of the 80's, the poverty-of-the-stimulus argument has been reformulated in rather extreme terms. Now it is claimed that central aspects of language are learned on the basis of not just limited data, but of no data at all. More precisely, it is claimed that, in order to learn his first language, the child needs to know that some forms are not correct; but since he has no 'negative' data that would directly tell him this, he must possess the requisite knowledge 'in advance'.[7] Thus, innate linguistic knowledge is invoked once again. - Let us see what is wrong with this argument.

Consider the following proportional analogy:

| | | |
|---|---|---|
| (1) Dad told a story to Sue | = | (3) Dad said something nice to Sue |
| (2) Dad told Sue a story | | (4) *Dad said Sue something nice |

The problem of 'negative evidence' is simply this: How does the child know that sentences like (4) are incorrect, without being explicitly told about their incorrectness? That is, how does he learn to block the 'false analogy' exemplified by the above sentences? Notice that this case, which involves the

coexistence of two correct structures (1) and (2), is at least on the face of it more difficult to explain than the standard type of morphological exception, for instance:

hand:hands :: foot:feet (*foots)

Here there is only one correct form in each case; and feet is a 'positive' exception, i.e. it occurs although it should not, whereas (4) is a 'negative' exception, i.e. it does not occur although it should (for more discussion, see Bowerman 1988 and Lasnik 1989).

The first thing to notice about this example is that it has nothing to do with innate principles of any sort. This is due to the fact that the variation between tell and say (or give and donate, or show and demonstrate) is too low-level to be anything else than just a matter of English language. Yet it is simply a fact that the child does learn to say (2) and to avoid (4). And since he learns cases like this without the aid of innate principles, there is no a priori reason to assume that he would need such principles in learning other cases (for instance, cases involving the Wh-movement).

There still remains a legitimate question concerning language-particular learning: how, precisely, does the child learn such idiosyncrasies of English like the variation between tell and say, or the quite subtle variations enumerated by Bowerman (1988: 90-93)? The detailed answer must be left to child-language experts, but at least the general answer is clear: It happens on the basis of positive evidence and without innate linguistic principles. Those who have doubted the child's ability to learn on the basis of positive evidence have simply underestimated the powers of innate analogical reasoning. It tells us not only which analogies to accept, but also which not to accept.

It should be clearly understood how unreasonable it was to assume, in the first place, that positive evidence is not enough. Having heard a sentence like Mary drinks milk, the child can form a sentence like John drinks beer. No one needs to tell him that he should not say John drinks to beer, or any such (incorrect) construction. It is enough that he never hears such a sentence.

Considering the importance of the poverty-of-the-stimulus argument to the generative paradigm, it is rather odd to notice that it is supported by no real data.[8] Claims are constantly justified by nothing but subjective impressions (more precisely, subjective impressions of sceptical nature), i.e. by claiming that it is 'not plausible' or 'hard to imagine' or 'difficult to believe' that the child can do such-and-such. But other linguists have found all this quite plausible, or easy to imagine.

Let us look at an example of this 'sceptical' type of argumentation, taken from Hoekstra & Kooij (1988: 37-38). Consider these two sentences:

(5) Where did John say that we had to get off the bus?
(6) Where did John ask whether we had to get off the bus?

(6) can be understood only in one way, whereas (5) has two interpretations, namely as asking either about the place where John said what he said or about the place where we had to get off the bus. Why do we know this difference? - according to Hoekstra and Kooij, because we have "access to complex principles of UG". Why are they needed? - because "this piece of knowledge[, although] shared by all native speakers,...can hardly have been established on the basis of induction [or analogy]". Why? - "simply because there are no data from which induction could conceivably proceed". But of course there are precisely such data, and self-evidently so. The first interpretation of (5) is understood in analogy with straightforward where-questions, for instance:

(7) John slept - Where did John sleep?

The second interpretation of (5) results from questioning an element in a that-clause, which means that it is understood in analogy with a sentence pair like the following:

(8) John said that Bill drank beer - What did John say that Bill drank?

The ambiguity of (5) can thus be explained by relating it either to (7) or to (8), each of which has only one interpretation. The analogical process involved in examples like this is formalized in Itkonen and Haukioja (forthcoming).

By this time we have come to see the circular nature of the poverty-of-the-stimulus argument (cf. also Campbell & Bickhard 1992: 581-582). The claim that children know linguistic facts for which they have 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. In sum, the stimulus that the child receives appears poor only on the assumption that he is incapable of handling it; and this incapacity is assumed to result from the fact that he is utterly devoid of any generalizing capacity. (Alternative terms are induction and analogy.)

3.2. Analogy

Analogical or inductive reasoning is today the object of intensive research (cf. Holland et al. 1986, Helman 1988, Vosniadou & Ortony 1989). Within generative linguistics, however, analogy is treated with suspicion. The reason for this attitude is Chomsky's long-standing hostility towards analogy, as expressed in the following statement:

"It seems that there is little hope in accounting for our [linguistic] knowledge in terms of such ideas as analogy, induction, association, reliable procedures, good reasons, and justification in any generally useful sense" (Chomsky 1986: 12; repeated on p. 55 and 222).

Chomsky's anti-analogical position can be shown to consist of several strands each of which is mistaken (cf. Itkonen & Haukioja forthcoming). In the present context I shall single out three of the most outstanding mistakes.

First, Chomsky formulates the issue as a contrast between (non-theoretical) item and (theoretical) grammar:

"It is easy to show that the new events that we accept and understand as sentences are not related to those with which we are familiar by any simple notion of formal (or semantic or statistical) similarity or identity of grammatical frame. Talk of generalization is entirely pointless and empty. It appears that we recognize a new item as a sentence not because it matches a familiar item in any simple way, but because it matches the grammar that each individual has somehow and in some form internalized" (Chomsky 1959: 56; emphasis added).

Later, e.g. in Chomsky (1965: 58), the target of criticism is defined more narrowly as 'physical similarity'. Thus, in conformity with his emphasis on linguistic form (cf. 2.4), Chomsky conceives of analogy as simple physical similarity between items. If the relations between sentences cannot be accounted for in such terms, he concludes that analogy fails, and that sentence-relations can only be accounted for implicitly, by a theoretical non-analogical grammar. This dichotomy is fallacious, however. The standard definition of analogy is structural (or relational) similarity, and not material (or

physical) similarity (cf. Hesse 1963: 67-68, Gentner 1989: 206-207). This means that, in language, analogy holds between sentences as abstract, grammatically analyzed entities, and not between sentences as physical, unanalyzed entities. This is the traditional view. For instance, Sapir (1921: 85) notes that the following two sentences "fit precisely the same pattern", "differing only in their material trappings" (emphasis added):

The farmer kills the duckling
The man takes the chick

This is how new words and new sentences "are being constantly created"..."on the analogy of old ones", i.e. "on strictly traditional lines" (p. 37). In exactly the same way, Jespersen (1965 [1924]: 19) claims that the following two sentences are 'analogous', or 'made after the same pattern':

John gave Mary the apple
My oncle lent the joiner five shillings

It is self-evident that in both cases the two sentences, although analogous, are physically quite dissimilar. This means that Chomsky is mistaken in thinking that 'analogy' equals '(simple) physical similarity'. Of course, he is quite correct to claim that the relations between sentences cannot be accounted for in terms of physical similarity. But he should be aware that this fact has no bearing at all on the viability of the notion of analogy.

Similarly, today when analogy remains an accepted part of diachronic linguistics, it is claimed that analogy "involves paradigmatic organization" and "makes the unobservable changes of reanalysis observable" (Hopper & Traugott 1993: 61). It should be clear that such concepts as 'paradigmatic organization' and 'reanalysis' which are used to define analogy are theoretical, i.e. not definable in 'simple physical' terms; and therefore analogy is an equally theoretical concept. To repeat, this is the traditional position. Chomsky's rejection of analogy is based on an arbitrary and unjustified redefinition. It is a mystery why this redefinition has been accepted with so little protest.

Second, analogy has traditionally been invoked to account for language-acquisition. It is obvious that language-acquisition is a gradual process, but Chomsky's anti-analogical position renders him uncapable to account for gradualness. It is only logical that he (e.g. 1986: 52) should make the 'simplifying idealization' that language-acquisition takes place instantaneously: it is only in this way that he can avoid admitting the fact that over the years the child constructs a series of grammars each of which is a (sometimes revisionist) extension or generalization of the previous one.[9]

Moreover, language-acquisition, which is a process, shades off into production and understanding of new sentences, which too are processes. Jackendoff (1987: 38-39) openly admits that Chomskyans have always been interested in structure, and not in process. Therefore, if we ask how processes (of production/understanding) should be described, it is inconsistent to point to descriptions of (syntactic) structures. That is, even if Chomsky were right in his rejection of analogy (which he is not), it would still be wrong (i.e. a category mistake) to claim that descriptions of structure give us descriptions of process. But precisely this was entailed by the original justification that Chomsky (1965: 57-58) offered for generative grammar, namely its alleged capacity to account for the "fundamental fact about the normal use of language, namely the speaker's ability to produce and understand instantly new sentences".[10]

Third, it is inconsistent to deny the existence of analogy, while practicing it at every turn. In Section 2.1 we already saw the analogical influence that artificial languages have exerted on the generative notion of language. Let us now consider the so-called clear case principle, which regulates the

generative attitude towards linguistic data insofar as it constitutes the basis for dividing the sentences into two discrete classes, viz. grammatical and ungrammatical:

"In many intermediate cases we shall be prepared to let the grammar itself decide, when the grammar is set up in the simplest way so that it includes the clear sentences and excludes the clear non-sentences. ... A certain number of clear cases, then, will provide us with a criterion of adequacy for any particular grammar" (Chomsky 1957: 14).

On closer inspection, the clear-case principle turns out to be based on analogy. The nature of sentences with an equivocal grammaticality status is decided in analogy with the unequivocally (un)grammatical sentences. That is, if an equivocal sentences is (structurally) more similar to a clearly grammatical sentence than to a clearly ungrammatical sentence, then it is decreed to be entirely grammatical; and vice versa.

The Arab grammarians were perceptive enough to realize that their central methodological concept was 'analogical extension' (= qiyaas) (cf. Itkonen 1991: 132-134). Because of Chomsky's hostility towards the notion of analogy, representatives of the generative paradigm have been prevented from achieving the same insight. Moreover, generativists practice analogical extension not just at the level of data, but also at the level of theoretical concepts, as we shall see in connection with our discussion of X-bar syntax, in Subsection 5.1.[11]

### 3.3. Learning forms without meanings?

It is evident from what precedes that Chomsky regards the primary linguistic data as syntactic (or formal or physical) in character. This entails the assumption that the primary type of language acquisition must be the learning of meaningless forms (cf. the Matthews-quotation above). Several arguments can be adduced to show that this assumption is mistaken.

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. It is hard to see why this fact should be ignored in the context of language-acquisition.

Second, when at the very first stage of language-acquisition the child lying in the crib hears meaningful words of his own lexicon and phonetically similar nonsense words, he repeats only the former (and does so excitedly). This shows that, to put it metaphorically, the child reaches out for meaning. Form is something contingent.

Third, humans have an innate capacity to endow human actions (and even natural events) with meaning. 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 sound symbolism. The same is true of hearing nonsense rhymes. The learning of pure form, if it ever occurs, is an abnormality.

Fourth, 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 one (optionally) deletes a subject in a 'pro-drop language', there is always a reason for doing so (e.g., 'because it was not needed'). The Chomskyan emphasis on 'autonomous syntax' (or pure form) requires us to envisage actions made for no reason at all.

Fifth, according to Chomsky's scenario, when a child hears a limited number of strings of sounds which may be identified 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. The fact is that this exposure does not bring about language-acquisition, unless it is accompanied by some explicit teaching. Thus mere sound (= 'pure form') is not enough. What is required, in addition, is the (natural) context of use, i.e. precisely that aspect which Chomsky is anxious to suppress.

Sixth, it is generally agreed today that spoken languages and sign languages stem from a common faculty. The pervasive iconicity of the sign languages (and in particular, of the pointing signs) makes it impossible 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.[12]

Seventh, let us have a closer look at the Matthews-quotation above. It is curious that Chomsky experiences insuperable difficulties in trying to figure out how the child manages to 'glean meaning from the context of utterance'. The associationist learning theory already provided part of the answer: When a child sees a dog and hears dog, he has learned that dog means 'dog'. Of course, associationism does not tell the whole story about language-acquisition, but it is silly to argue that just because it is not enough, it does not exist. It is certainly no accident that children's earliest vocabulary contains only items with whose referents they have been directly acquainted. For Chomskyans, however, this remains an impenetrable mystery.[13] 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).[14]

Finally, it has to be asked if there is any evidence at all for the notion of 'formal learning'. Chomsky (1992: 1) refers to the occurrence of a syntax-semantic dissociation in some mentally retarded persons (cf. Yamada 1990, Smith & Tsimpli 1991). However, this phenomenon is rare enough to be comparable to the extraordinary capacities of so-called 'idiots-savants' (cf. Haukioja 1993). Because the learning of e.g. mathematics is not explained by invoking the 'idiots-savants', there is no reason why language-learning should, or could, be so explained. It may be added that the 'savant' linguist investigated by Smith & Tsimpli (1991) does not exhibit any (systematic) syntax-semantics dissociation.

Thus I conclude that formal learning does not exist. In view of this, it is interesting to note that there is a vast literature on so-called learnability theory, which is based precisely on the assumption of formal learning (cf. Matthews & Demopoulos 1989). Within this framework, the input data (also called 'learning environment') consists of syntactic entities, i.e. sentences or surface strings, which serve as the basis for hypothesizing a grammar that might generate them. It has sometimes been suggested that even if learnability theory rests on a false assumption, the precision that it brings to the study of language-acquisition is valuable in itself. My reply to this is that if someone claims to be able to measure the flatness of the Earth with nanometric precision, I am not impressed (because the Earth is not flat). I much prefer a less precise but more truthful description like "The Earth is roughly spherical".

What about the learning of word-meanings? This is the state-of-the-art view of the matter:

"Even after relational terms have entered the vocabulary, children are slow to acquire their full meanings...The correct usage of common verbs...is not fully mastered until rather late (5 or 6 years of age and, in some cases, as late as 8 years or older)" (Gentner & Ratterman 1991: 254).

It is quite instructive to compare this account with Chomsky's opinion:

"[The meaning of the word climb] is very complicated. But every child learns it perfectly right away. Now that can only mean one thing. Namely, human nature gives us the concept 'climb' for free.That is, the concept 'climb' is just part of the way in which we are able to interpret experience available to us before we even have the experience. That is probably true for most concepts that have words for them in language. This is the way we learn language" (Chomsky 1988: 190-191; emphasis added).

The latter quotation reveals the startling fact that Chomsky apparently has no idea whatever of how the child really acquires his native language. He seems honestly to believe that having heard a very limited number of partly mispronounced sentences, the child learns the language "perfectly right away". No wonder that he thinks the child needs the aid of an innate grammar. In the present context the only interesting question is why someone with no interest in language-acquisition should have adopted its explanation as his ultimate goal. This question will be answered in Section 6.

## 4. Universal grammar

### 4.1. Principles and parameters

Chomsky's universal grammar (henceforth to be abbreviated as 'PP') postulates the existence of innate principles and parameters. Principles are assumed to be valid in all languages. For instance, the principle of structure-dependency says that linguistic operations are performed on (hierarchic) structures, rather than atomary units. To give another example, the binding principles define a 'local domain', and state that anaphors are bound (i.e. must have an antecedent) in their local domain, whereas pronominals are free. By contrast, parameters admit of inter-language variation. For instance, English chooses the value '+' on the parameter of Wh-movement, whereas Japanese and Kwakwala choose the value '-', because they have no ('visible') Wh-movement.[15] 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 a (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 respect to the other material (i.e. specifiers and complements) contained in the phrases. The parameter of primary branching direction, according to which sentences are embedded either to left or to right, is an elaboration of the head parameter. Suffixing vs. prefixing constitutes a parameter of morphological order. Finally, any language must choose either '+' or '-' on the pro-drop parameter, i.e. it may or may not suppress the subject of a clause. The number of the parameter values is limited, preferably to only two.

Structure-dependency is an unrestricted universal. Subjacency, by contrast, is a restricted or implicational 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.

The principles and parameters are thought to constitute the 'core' of any language; everything else is 'periphery'. The core is unmarked whereas the periphery is marked. There is markedness also within the core, because one of the two values of a parameter is unmarked while the other is marked.

Any purportedly universal grammar is accountable for the data on which it bases its (universalist) claims. This topic will be discussed in the next two subsections. At the same time, I cannot help touching upon some issues that will be treated more extensively in Section 5.

4.2. '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 1980a: 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 1991: 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, disregarding any surface variation among mosquitos, I shall make the bold prediction that all animals fly and have the size of less than 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 <u>how probable</u> it is that universalist claims based on one language, or on fortune-cookies, or on christal balls turn out to be <u>true</u>. And the answer is that in all three cases it is about equally <u>improbable</u>. 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 (types of) languages as possible.

At first glance, it might seem that 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 '-'.) Nevertheless, the basic correctness of the one-language approach is still stoutly defended:

"Although the investigation [in the previous chapters] has been restricted to one language, some of the principles isolated and tested for validity are sufficiently deep and general to allow us to make a preliminary claim as to their universal character" (Ouhalla 1994: 268).

In the next subsection we shall see that these words are meant to be taken seriously.

### 4.3. PP and cross-linguistic evidence

With the 'principles-and-parameters' approach Chomsky's universal grammar seems to have opened itself to cross-linguistic evidence. However, this change of attitude is more apparent than real.

Let us consider the lexical categories of the X-bar syntax, i.e. N, V, A, and P. On the basis of cross-linguistic evidence, Schachter (1985: 6) is forced to conclude that there are no universally valid formal criteria for defining the major lexical categories. Because the generative paradigm rejects the use of semantic (or ontological) criteria, it follows that its lexical categories 'hang in the air'.

Cross-linguistic evidence shows, moreover, that N, V, A, and P are not at all on an equal footing. N and V are open classes, and they are present in every language. A is an 'intermediate' category in the sense that it is absent in some languages, and in several languages it is a closed class (cf. Croft 1991: 130-133). P is absent in most Australian languages (cf. Dixon 1980: 272). A West African language like Yoruba has only two genuine instances of P, which may then combine with a small number of nouns to form preposition-like expressions (cf. Rowlands 1969: 29, 139-141). Kwakwala has at most two instances of P (Anderson 1992: 30). Thus, P is 'transitory', rather than intermediate, because it generally develops out of other categories (cf. Croft 1991: 144-146). For instance, components of Chinese serial verbs have become prepositions in the course of some 3'000 years (cf. Li & Thompson 1974). The same process has started more recently in Yoruba, both in serial verbs (cf. Givón 1976: 82-86) and in the P+N constructions just mentioned.

If one takes cross-linguistic evidence seriously, N, V, A, and P should not be treated alike (and should be provided with semantic definitions). So why are they treated alike in PP? Precisely because PP does not take cross-linguistic evidence seriously. The four lexical categories seem well-established in English. Apparently it is still assumed (although no longer openly declared) that what is true of English, must be true of Universal Grammar.

Let us consider another example. 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 (or so it seemed). 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, isolating languages like Chinese or Yoruba 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'. The case-systems of all isolating languages just happen to be morphologically non-realized. - This is one more application of the 'depth vs. surface' distinction as it was practised in the 60s: The facts are non-uniform; thus, postulate a level where everything is uniform, and call it 'depth'; call the facts 'surface', and forget about them.

Let us consider one more example. Greenberg (1966) noted certain less than perfect correlations (or 'tendencies') between the word orders within different types of phrases; 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 head parameter 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. 4.1 above). 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.[16]

Finally, consider the notion of VP in Universal Grammar. The generative paradigm assumes that V and its complement NP must be adjacent. A VSO language like Welsh or Kwakwala shows that this is not always the case. Instead of accepting the cross-linguistic evidence, however, generativists concentrate their attention on "trying to work out how the VSO order is derived from a structure where the verb and its complements form a VP and are adjacent at D-structure" (Ouhalla 1994: 290). The inspection of cross-linguistic evidence turns out to be a distraction, because insofar as it differs from the model provided by English, it has to be reduced to the latter.

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 non-European languages do come under scrutiny, they are either reduced to the model of English or they just have the function of providing marked values for parameters (cf. also 5.1). The de facto ornamental role of cross-linguistic evidence shows that the one-language approach is still lurking in the background.

## 5. Methodological notions

### 5.1. Falsification

It is one of the putative virtues of PP (as well as of its predecessors) that, being highly formalized, it makes very specific claims: because of their specificity, it should be possible to see immediately whether the claims are 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 is fallacious. In the present subsection I shall explain in somewhat more detail why this is so.

The degree of formalization is irrelevant to falsification, if the relation between the description and the data remains uncertain. Let us consider the X-bar syntax. It is based on the assumption that all lexical categories share the same structure, namely specifier, adjunct, head, and complement. The head is obligatory, and whether or not it has a complement, is stipulated in the lexicon. The specifier and the adjunct are optional (cf. Ouhalla 1994: Chap. 4). (To be sure, we have already seen in Subsection 4.3. that it is wrong, in the first place, to treat the heads N, V, A, and P alike.) Because the order of the constituents is in practice free (i.e. the exceptions are merely labelled as 'marked'), it is hard to see how this general conception could be falsified.

Next, the structure of the lexical categories is generalized to the structure of the non-lexical category S. This generalization is made possible by taking S to have the tripartite structure NP AUX VP, where AUX functions as the head. AUX is reinterpreted as I (= 'Inflection'), which stands for Tense and Agreement. Thus, S(entence) turns out to be IP (= 'Inflection Phrase').

AUX is a lexical category. In English the postulation of AUX is supported by the temporal/modal verbs will, can, etc. and by the do-support, whereas the auxiliary verbs proper be and have are taken to originate within VP. (In languages like Finnish or Swahili there is no comparable reason to postulate AUX as a category separate from V; cf. Perrott 1951: 128.) As we just saw, AUX is reinterpreted as the functional category I, and S is reinterpreted as IP. This reinterpretation is far from natural because the specifier of IP, i.e. the subject NP, differs from the specifiers of other categories in being obligatory, and a separate rule, called 'Extended Projection Principle', is needed to stipulate this. Thus, the real reason for postulating IP turns out to be the compulsion exerted by 'theoretical analogy': it is so 'neat' to assume that all categories, both lexical and non-lexical, have the same structure.

In what precedes, the structure of S was modelled on that of (e.g.) NP. There is an interesting sequel to this maneuver. If the head of IP is a functional category like I, how can it be that the head of NP is a non-functional category like N? Here theoretical analogy sees a new opening: In reality, the head of NP is not N but the determiner (e.g. the definite article the), which means that NP must be reinterpreted as DP (= 'Determiner Phrase)' (cf. Ouhalla 1994: 179-182). Thus, the structure of IP (formerly S) is generalized to that of NP (henceforth DP), after the structure of NP had first been generalized to S. This series of (re)definitions has no observable consequences.

The foregoing account was simplified insofar as theoretical analogy had already been at work in the definition of S: because the subordinate clauses are generally preceded by a complement (in English that, for instance), and because it is neat to assume that the subordinate clauses and the main clauses share the same structure, it seems 'natural' to make the generalization that the main clauses too are preceded by a complement (cf. Ouhalla 1994: 62-65). Never mind that it is invisible (except in questions, in some languages). Notice how difficult it is to deny that something invisible exists.

Such breath-taking reinterpretations and generalizations can only be performed by people who are obsessed with theoretical analogy. In view of this, it is rather ironical that of all people, the representatives of the generative paradigm should deny the existence of analogy (cf. Subsection 3.2). More importantly, there is no obvious way (and perhaps, no way) that the results of their theoretical analogy, although formalized with extreme precision, could be amenable to falsification.

Thus, a sufficient distance between the description and the data is a general guarantee against falsification. In the 60s Chomsky made also use of two more specific '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).[17] Insofar as Greenberg-type universals do not agree with PP, they are declared to be part of the periphery. The 'core vs. periphery' distinction is a generalization of the older 'competence vs. performance' distinction, as can be seen by comparing the quotation from Cook (1988) with this somewhat older quotation:

"An acceptable theory of the relation between competence and performance models will have to represent that relation as abstract, the degree of abstractness being proportional to the failure of formal features of derivations to correspond to performance variables" (Fodor & Garrett 1966: 152).

These two quotations, which are separated by more than twenty years, reveal the essence of the generative cast of mind. If the facts seem to contradict the theory, either the facts are wrong (= 'marked') or the distance between the theory and the facts has to be increased (or made 'more abstract'). The possibility that there might be something wrong with the theory transcends the average generativist's intellectual horizon.

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 and prepositions. 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'. For instance, the analysis of serial verbs in Yoruba forces Baker (1988) to assume that VP can have two heads. This entitles him to set up a VP parameter ('one head vs. two heads') where Yoruba represents the marked value.

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 merely reinventing the notion of (Greenbergian) implicational universal.

The immunization strategies which I have spoken about here are combined with 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 data of any kind. 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 generative 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, generative 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.[18] But when subsequent experiments invalidated the assumption of deep structures and transformations, he 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 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).[19] 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. It is far from clear that there is indeed some systematic justification for the changes that the succession 'TG > GB > PP > minimalism' exhibits. Rather, the changes seem to have been 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 can influence, the development that started with TG. (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.

5.2. 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:

"There is no a priori reason why human language should make use exclusively of structure-dependent operations, such as English interrogations, instead of structure-independent operations [like the operation that inverts the first and last words of a sentence]" (Chomsky 1968: 52).

"...the universals that rationalists typically cite (the structure-dependency, the subjacency condition, etc.) are certainly unexpected and nonintuitive" (Matthews 1989: 69).

The possibility of any functional explanations for Chomskyan universals has been ruled out explicitly:

"To account for or somehow explain the structure of UG, or of particular grammars, on the basis of functional considerations is a pretty hopeless prospect, I would think; it is, perhaps, even 'perverse' to assume otherwise" (Chomsky 1975b: 58).

As these quotations make clear, innateness lends itself quite naturally to being some sort of deus ex machina: when you do not know what to say about something, say that it is innate. To give another example, when Katz (1981) defends his Platonism, he is confronted with the dilemma noted already by Aristotle: how is it possible that we, beings who live in space and time, come to know something that is beyond space and time, i.e. Platonic ideas? As you might have guessed, Katz's answer is that we are just innately equipped for it (cf. Itkonen 1983b). Similarly, when John E. Joseph tried to account for the normativity of language at the World Congress of Linguists in Québéc, August 1992, he found no other way than to postulate separate innate modules for normativity and for language (which is like postulating separate modules for the sides of a rectangle).

It has often been pointed out that the appeal to innate ideas 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, generativists have strongly rejected this interpretation, but their reasons for doing so remain confused. Hoekstra & Kooij (1988), for instance, refer to the 'theoretical foundation' that generativists presumably possess and their opponents lack: whether or not a universal principle is decreed to be innate, is claimed to result 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 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; cf. Deane 1991).

The conceptual confusion that prevails 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 this 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; emhasis 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 worthwhile to start looking for them. But of course, the evidence for innateness is far from overwhelming. In fact, the existence of pervasive morpho-syntactic iconicity proves that linguistic structure is largely, and perhaps exhaustively, explainable (for an overview, see Itkonen 1994). In the present context I give only the outline of this argument.

The explanation of structure-dependency, for instance, 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.[20] - 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/cognition' (as he should).

Sometimes it is said that even if a principle like subjacency cannot be explained (as Chomskyans see it), it explains something, namely why NPs can be moved in some ways, and 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 ungrammaticalness 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 shown 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'), and I may even predict that any other triangles that will perhaps appear will also be red. However, it would be inappropriate to say that I have explained anything. In particular, I have not explained why this thing is red, if I have stated that it is a triangle. A genuine eplanation 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 in the way we do speak. But this is the 'virtus dormitiva' strategy. In just the same way Katz (1964) 'explained' 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 /Iz/, because in their heads they have the mechanism which makes them form the plurals in the way they do, namely by adding the morpheme {S} with the three allomorphs /s/, /z/, and /Iz/ (cf. Itkonen 1978: 200-202).

We seem to have reached the following conclusion: Chomsky's universal grammar represents a level of autonomous syntax and, being innate, it cannot be explained by any functional considerations. Interestingly enough, this view has recently been called into question also within the generative paradigm. Jackendoff (1992), although claiming allegiance to the autonomy of syntax (p. 31), nevertheless admits that because "syntax presumably evolved as a means to express conceptual structure, it is natural to expect that some of the structural properties of concepts would be mirrored in the organization of syntax" (p. 39). Thus, instead of being autonomous and innate, (at least part of) syntactic structure is causally explained, namely as the expression of (prelinguistic) conceptual structure. Similarly, having noted that spatial language makes very fine distinctions between the physical shapes of objects, but is much less constrained when it has to express physical locations and movements, Jackendoff (1992: Chapter 6) refuses to accept this fact as just an aspect of the innate linguistic endowment. Rather, he wishes to explain it by postulating a distinction between 'what' and 'where' in the organization of spatial representation, and by regarding the linguistic asymmetry merely as a reflection of this conceptual asymmetry. Inconsistently, however, he thinks he has shown that in this area there is no need for functional explanations (e.g. explanations referring to the efficacy of the asymmetry in question). He fails to see that his own explanation is thoroughly functional, though in a more general sense: It is the function of language to speak of the external world as it has been conceptualized by man.

In a more direct fashion, Newmeyer (1990 and 1991) frankly admits that there are quite plausible functional explanations for phenomena like structure-dependency and subjacency. Because this explicitly contradicts what Chomskyans have claimed earlier (cf. above), it certainly looks like a falsification of the Chomsky-type innate universal grammar. Astonishingly, Newmeyer denies this. As he sees it, the functions that language may be shown to serve have promoted the survival of the human species; therefore they have become, via mutations, part of our innate language capacity; therefore, rather than falsifying Chomsky's innate universal grammar, they actually confirm it. - After reading this remarkable argument, I became convinced that there is nothing, absolutely nothing, that could make Chomskyans admit that there is or has ever been anything amiss with their theory. In the game of linguistics, truth is a secondary consideration; not to lose face has top priority.

## 6. Conclusion: Chomskyan linguistics is an explanans in search of an explanandum

Once the reader has reached this point, he may have started to puzzle over the nature of Chomskyan linguistics: it is a psychological or even biological theory without any psychological or biological facts, a theory looking for universal features, but telling in advance that whatever it will find, is unexplainable. In my opinion, these puzzling aspects of Chomskyan linguistics can be understood only if they are put in an historical perspective.

In his 1955 dissertation Chomsky was still a follower of Bloomfield and Harris. He defended the former's antimentalist program against suggestions that the criteria of scientific significance should be relaxed so as to admit mental and non-formal entities as well:

"The fact that a certain general criterion of significance has been abandoned does not mean that the bars are down and that 'ideas' and 'meanings' become proper terms for linguistics. If this rejection of an old criterion...is followed by a new analysis of 'significance', then if this is at all adequate, it seems to me that it will rule out mentalism for what were essentially Bloomfield's reasons, i.e., its obscurity and inherent untestability" (1975a [1955]: 86).

It was Chomsky's goal to achieve a synthesis of Bloomfield's antimentalism and Harris' distributionalism at the level of syntax:

"The notions that enter into linguistic theory are those concerned with the physical properties of utterances, the formal arrangement of parts of utterances,...and finally, formal properties of systems of representation and of grammars...We will refer to linguistic analysis carried out in these terms as 'distributional analysis'" (p. 127), "this term [being] borrowed from Harris" (p. 63, n. 1).

Chomsky (1957: 55) also echoes Harris (1946: 164, n.6) in his claim that "the only ultimate criterion in evaluation is the simplicity of the whole system". So we see that, contrary to a wide-spread misconception, in its first explicit formulation Chomsky's linguistic theory did not embody any significant break with the preceding tradition. On this issue, the level of historiographical writing, epitomized by Newmeyer (1980), has been extremely low.

In Chomsky (1965) the study of linguistic form, now supplied with a mentalist interpretation, was seen as a study of innate universal grammar. Since Chomsky (1980b) the innate linguistic 'organ' has been conceived of in ultimately biological terms. The autonomy of this organ vis-à-vis other comparable organs is justified by a general appeal to the modularity of the human psycho-biological constitution. If the mind is taken to rest on a neurophysiological foundation, and if language is declared to be a module of the mind, and if syntax is declared to be the central module of language, then it might almost seem that one has indeed acquired the right to study English syntax and call it neurophysiology.

In the present context it is of no importance that Chomsky's theory of syntax has undergone several modifications. What is important, is the fact that while he has continued to analyze the syntax of English by means of self-invented sentences which his own linguistic intuition deems either correct or incorrect, his interpretation of, and justification for, what he is doing has changed completely: from antimentalist distributional analysis he has moved first to mentalist syntax and then to biology.

Once the generative syntax had been invented, something had to be done with it, i.e. it had to be used to 'explain' something. With the passing of time the explanandum has been conceived of in increasingly ambitious terms: having started with distributional arrangements of English morphemes, Chomsky has now arrived at theoretical biology. Seen in perspective, innatism and modularity are not claims with empirical content. They are just excuses for Chomsky not to do anything different from what he has always done.

It is only human that Chomsky should have wished to squeeze every possible, and even impossible, gain out of his great invention, i.e. generative syntax. The criticism really directs itself against his disciples who, forsaking their duty of critical thought, have let this happen.

Notes

1.Chomsky adds that "they played virtually no role in early work on generative grammar except in informal exposition, or since". Compare this withe following quotation: "The fundamental aim in the linguistic analysis of a language L is to separate the *grammatical* sequences which are the sentences of L from the *ungrammatical* sequences which are not sentences of L and to study the structure of the grammatical sequences...we assume intuitive knowledge of the grammatical sentences of English ..." (Chomsky 1957: 13; emphasis in the original). It is also good to recall that Chapter V of Chomsky (1975a [1955]) bears the title 'Grammaticalness'.

2. Written documents, for instance, do not get me out of this circle, because now the question arises whether I remember correctly the meanings of the written 'private' words. Kenny (1973: 192-193) presents this argument exceptionally well. Saunders & Henze (1967) offer a more exhaustive discussion.

3. In practice, of course, subjective (linguistic) intuition produces reliable results most of the time.

4. It may be added that in this context Wittgenstein's originality has been overrated. Hegel already presented the same argument, in practically identical terms.

5. It is good to point out that Lakoff-type cognitive semantics is anxious to find a 'grounding' in physical and social experience, in a more or less Wittgensteinian spirit.

6. The postulation of '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: Chapter 3.

7. Notice, however, that slips of the tongue are typically *corrected*, which could be taken to mean that incorrect forms do come labelled as incorrect; cf. Wittgenstein (1958, I, § 54): "But how does the observer distinguish...between players' mistakes and correct play? - There are characteristic signs of it in the players' behavior. Think of the behavior characteristic of correcting a slip of the tongue. It would be possible to recognize that someone was doing so even without knowing his language."

8. For instance, Chomskyans have been trying to explain the 'fact' that children do not perform syntactic overgeneralizations. But Bowerman (1988) shows that there is no such fact.

9. Another way to ignore the gradual nature of language-acquisition is to view the 'continuity' between successive child grammars merely as a matter of 'fixing the parameters' (cf. Hyams 1987). Even assuming, for the sake of argument, that 'parameter' is a viable notion, this stance omits a huge number of facts, in particular the *growth* evinced by successive grammars.

10. This ability was called the "creative aspect of language use" (Chomsky 1965: 6). Initially, creativity was identified with recursivity: "recursive rules...provide the basis for the

creative aspect of language use" (Chomsky 1967: 7). This was soon admitted to have been a mistake: "What I have elsewhere called the 'creative aspect of language use' [cannot] be identified with the recursive property of grammars. [M y f]ailure to keep these very different concepts separate has led to much confusion" (Chomsky 1975b: 230, n. 11).

11. The limits of the clear case principle are explored in Itkonen (1976).

12. Some people would like to contest the iconic nature of sign languages. However, even if particular sign are (or have become) non-iconic, the very structure of any given sign language is based on the (iconic) idea of modelling the reality which is spoken about. That is, the space in front of a signer is a miniature model of the world, and 'place-holders' for real-life entities are first put in it, and then pointed at and moved around in accordance with the exigencies of the story to be told. Those who deny the iconic nature of sign languages seem to assume that admitting it would commit them to admitting that sign languages are not genuine languages on a par with spoken languages. This is a mistake, however. Spoken languages too are iconic in character.

13. Cf. Fodor (1983: 81): "It is a standing mystery in psychology why...subjects should exhibit a reliable and robust disposition to associate 'salt' with 'pepper', 'cat' with 'dog', 'mother' with 'father', and so forth."

14. 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, regardless of what is true, could be called an instance of the 'argument from laziness'.

15. I disregard the possibility of 'logical' or 'invisible' movement.

16. 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.

17. 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*'.

18. 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."

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

20. Notice that even a sentence like *The woman was planning a party* reflects the fact that I 'see' the woman together with the party, although in distinct 'mental spaces'.

References

Anderson, Stephen R., 1992. A-morphous morphology. Cambridge: Cambridge University Press.
Baker, Mark C., Object sharing and projection in serial verb constructions. Linguistic Inquiry 20: 513-533.
Bickerton, Derek. 1981. Roots of language. Ann Arbor, Mich.: Karoma.
Bloomfield, Leonard, 1936. Language or ideas? Language 12: 89-95.
Bowerman, Melissa, 1988. The 'no negative evidence' problem. In: Hawkins, ed., 1988: 73-101.
Butterworth, George and Lesley Grower, 1988. The origins of referential communication in human infancy. In: Lawrence Weiskrantz, ed., Thought without language, 5-24. Oxford: Clarendon Press.
Campbell, Robert L. and Mark H. Bickhard, 1992. Clearing the ground: Foundational questions once again. Journal of Pragmatics 17: 557-602.
Carnap, Rudolf, 1937. The logical syntax of language. London: Routledge.
Chomsky, Noam, 1957. Syntactic structures. The Hague: Mouton.
Chomsky, Noam, 1959. A review of B.F. Skinner's "Verbal behavior". Language 35: 26-58.
Chomsky, Noam, 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.
Chomsky, Noam, 1967. Recent contributions to the theory of innate ideas. Synthese 17: 2-11.
Chomsky, Noam, 1968. Language and mind. New York: Harcourt.
Chomsky, Noam, 1969. Quine's empirical assumptions. In: Donald Davidson and Jaakko Hintikka, eds., Words and objections. Essays on the work of W.V. Quine, 53-68. Dordrecht: Reidel.
Chomsky, Noam, 1975a [1955]. The logical structure of linguistic theory. New York: Plenum Press.
Chomsky, Noam, 1975b. Reflections on language. New York: Pantheon.
Chomsky, Noam, 1980a. On cognitive structures and their development. In: M. Piattelli-Palmarini, ed., Language and learning: The debate between Jean Piaget and Noam Chomsky. London: Routledge.
Chomsky, Noam, 1980b. Rules and representations. New York: Columbia University Press.
Chomsky, Noam, 1986. Knowledge of language. New York: Praeger.
Chomsky, Noam, 1988. Language and problems of knowledge. Cambridge, Mass.: MIT Press.
Chomsky, Noam, 1992. A minimalist program for linguistic theory. MIT Occasional Papers in Linguistics, 1.
Comrie, Bernard, 1981. Language universals and linguistic typology. Oxford: Blackwell.
Cook, V.J., 1988. Chomsky's universal grammar. London: Blackwell.
Croft, William, 1990. Typology and universals. Cambridge: University Press.
Croft, William, 1991. Syntactic categories and grammatical relations. Chicago: The University of Chicago Press.
Deane, Paul, 1991. Limits to attention: A cognitive theory of island phenomena. Cognitive Linguistics 2: 1-63.
Dixon, R.M.W. 1980. The languages of Australia. Cambridge: Cambridge University Press.
Fodor, Jerry A., 1975. The language of thought. New York: Crowell.
Fodor, Jerry A., 1983. The modularity of mind. Cambridge, Mass.: MIT Press.
Fodor, Jerry A. and Merrill F. Garrett, 1966. Some reflections on competence and performance. John Lyons and Roger Wales (eds.): Psycholinguistic papers. Proceedings of the 1966 Edinburgh conference. Edinburgh: Edinburgh University Press.
Gelman, Susan and James Byrnes, eds., 1991. Perspectives on language and thought. Cambridge: University Press.
Gentner, Dedre, 1989. The mechanisms of analogical learning. In: Stella Vosniadou and Andrew Ortony, eds., 1989: 199-241.
Gentner, Dedre and Mary Jo Ratterman, 1991. Language and the career of similarity. In: S. Gelman and J. Byrnes, eds., 1991: 225-277.
Givón, Talmy, 1975. Serial verbs and syntactic change: Niger-Congo. Charles N. Li, ed.: Word order and word order change, 47-112. Austin: University of Texas Press.
Goldin-Meadow, Susan and Carolyn Mylander, 1990. Beyond the input given: The child's role in the acquisition of language. Language 66: 323-355.
Goodluck, Helen, 1986. Language acquisition and linguistic theory. In: P. Fletcher and M. Garman, eds., Language acquisition. Cambridge: University Press.
Greenberg, Joseph, 1966. Some universals of grammar with particular reference to the order of meaningful elements. In: J. Greenberg, ed., Universals of language, 2nd ed., 73-113. Cambridge, Mass.: MIT Press.
Harris, Zellig, 1946. From morpheme to utterance. Language 22: 161-183.
Haukioja, Timo, 1993. Hyperlinguistic individuals and the relation between language and cognition: A methodological note. Cognitive Linguistics 4: 395-398.
Hawkins, John, 1985. Complementary approaches in universal grammar. Language 61: 569-587.
Hawkins, John, ed., 1988. Explaining language universals. London: Blackwell.
Helman, David H., ed., 1988. Analogical reasoning. Dordrecht: Kluwer.

Hesse, Mary, 1963. Models and analogies in science. London: Sheed & Ward.
Hoekstra, Teun and Jan Kooij, 1988. The innateness hypothesis. In: J. Hawkins, ed., 1988: 31-55.
Holland, John H., Keith J. Holyoak, Richard E. Nisbett, and Paul R. Thagard, 1986. Induction: Processes of inference, learning, and discovery. Cambridge, Mass.: MIT Press.
Hopper, Paul and Elisabeth Traugott, 1993. Grammaticalization. Cambridge: University Press.
Husserl, Edmund, 1913. Logische Untersuchungen, 2nd ed. Tübingen: Niemeyer.
Hyams, Nina, 1987. Parameters and syntactic development. In: Thomas Roeper and Edwin Williams, eds., Parameter setting, 1-22. Dordrecht: Reidel.
Itkonen, Esa 1976. The use and misuse of the principle of axiomatics in linguistics. Lingua 38: 185-220.
Itkonen, Esa, 1978. Grammatical theory and metascience. Amsterdam: Benjamins.
Itkonen, Esa, 1983a. Causality in linguistic theory. London: Croom Helm.
Itkonen, Esa, 1983b. A review of J. Katz's "Language and other abstract objects". Lingua 60: 238-244.
Itkonen, Esa, 1991. Universal history of linguistics: India, China, Arabia, Europe. Amsterdam: Benjamins.
Itkonen, Esa, 1994. Iconicity, analogy, and universal grammar. Journal of Pragmatics 22: 37-53.
Itkonen, Esa and Jussi Haukioja, forthcoming. Analogy in synchronic syntax (and elsewhere).
Jackendoff, Ray C., 1987. Consciousness and the computational mind. Cambridge, Mass.: MIT Press.
Jackendoff, Ray C., 1992. Languages of the mind. Cambridge, Mass.: MIT Press.
Jespersen, Otto, 1965 [1924]. The philosophy of grammar. New York: Norton.
Katz, Jerrold J., 1964. Mentalism in linguistics. Language 40: 124-137.
Katz, Jerrold J., 1981. Language and other abstract objects. Oxford: Blackwell.
Kenny, Anthony, 1975. Wittgenstein. Harmondsworth: Penguin Books.
Kripke, Saul, 1982. Wittgenstein on rules and private language. Cambridge, Mass.: Harvard University Press.
Labov, William, 1972. Sociolinguistic patterns. Philadelphia: University of Pennsylvania Press.
Lasnik, Harold, 1989. On certain substitutes for negative data. In: R Matthews and W. Demopoulos, eds., 1989: 89-105.
Li, Charles N. and Sandra A. Thompson. 1974. Historical change of word order: a case study of Chinese and its implications. John M. Anderson and Charles Jones (eds.): Historical linguistics I: Syntax, morphology, internal and comparative reconstruction, 200-217. Amsterdam: North Holland.
Markman, Ellen, 1991. The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In: S. Gelman and J. Byrnes, eds., 1991: 72-106.
Matthews, Robert J., 1989. The plausibility of rationalism. In: R. Matthews and W. Demopoulos, eds., 1989: 51-75.
Matthews, Robert J. and William Demopoulos, eds., 1989. Learnability and linguistic theory. Dordrecht: Kluwer.
Newmeyer, Frederick, 1980. Linguistic theory in America. New York: Academic Press.
Newmeyer, Frederick, 1990. A review of J. Hawkins' "Explaining language universals". Journal of Linguistics 26:
Newmeyer, Frederick, 1991. Functional explanation in linguistics and the origins of language. Language and communication 11: 3-28.
Ouhalla, Jamal, 1994. Transformational grammar. London: Arnold.
Perrott, D.V. 1951. Swahili. London: Hodder & Stoughton.
Pinker, Stephen, 1994. The language instinct. New York: Morrow.
Popper, Karl, 1963. Conjectures and refutations. London: Routledge.
Ravin, Yael., 1990. Lexical semantics without thematic roles. Oxford: University Press.
Rowlands, E.C. 1969. Yoruba. London: Hodder & Stoughton.
Sampson, Geoffrey, 1980. Making sense. Oxford: University Press.
Saunders, John and Donald F. Henze, 1967. The private-language problem. New York: Random House.
Schachter, Paul, 1985. Parts-of-speech systems. Timothy Shopen (ed.): Language typology and syntactic description, vol. I., 3-61. Cambridge: Cambridge University Press.
Schlesinger, I., 1975. Grammatical development: The first steps. In: E.H. Lenneberg and E. Lenneberg, eds., Foundations of language development, vol. I, 203-222. New York: Academic Press.
Slobin, Dan I., 1973. Cognitive prerequisites for the development of grammar. In: C. Ferguson and D. Slobin, eds., Studies of child language development, 45-54. New York: Springer.
Smith, Neil and Ianthi M. Tsimpli, 1991. Linguistic modularity? The case study of a 'savant' linguist. Lingua 84: 315-351.
Vosniadou, Stella and Andrew Ortony, eds., 1989. Similarity and analogical reasoning. Cambridge: University Press.
Wald, Benji, 1990. Swahili and the Bantu languages. Bernard Comrie (ed.): The major languages of South Asia, The Middle East, and Africa, 285-308. London: Routledge.
Wittgenstein, Ludwig, 1958. Philosophical investigations. Oxford: Blackwell.
Yamada, J. E., 1990. Laura: A case study for modularity of language. Cambridge, Mass.: MIT Press.