4

LANGUAGE AS A MENTAL FACULTY: CHOMSKY'S PROGRESS

P.H. MATTHEWS

Noam Chomsky is at once a brilliant grammarian and an important philosopher of language. As a grammarian, he has had greater influence on our conception of English syntax, both of the nature of syntax and the nature of particular constructions, than any other scholar now living, and continues to display a remarkable ability to discover new problems and new generalisations that his predecessors had entirely failed to notice. As a philosopher of language, he is responsible above all for the belief that linguistics is, in his terms, a branch of cognitive psychology, and that human beings have a genetically inherited faculty of language which is independent of other faculties of the mind. If these contributions were separate, they might well be thought to merit two chapters in an encyclopaedia of this kind. But they are intimately related. Chomsky's philosophy of mind rests directly on a philosophy of grammar, in which the term 'grammar' was used, in the 1960s, to refer not simply to a linguist's description of a language, but to the basic knowledge of linguistic structures that every speaker of a language has acquired in infancy. The central issues of linguistic theory are then posed as follows. First, we must ask what grammars are like: what form does a speaker's basic knowledge of a language take? Second, we have to ask how speakers do in fact acquire this knowledge. Chomsky's answer to the second question largely reflects his answer to the first, and both are central to his view of mind in general. The term 'philosophy of grammar' will recall the title of a famous work by Otto Jespersen (1924), a scholar with whose interests Chomsky has himself expressed sympathy (1975, 1986: 21f.). The aim of this chapter is to examine the development of his own philosophy of grammar, from its beginning in the 1950s to the form in which we find it now, thirty years after the work which first made his reputation.

I have referred, in the singular, to Chomsky's 'philosophy' of grammar. Like that of any other major scholar, his work forms a historical unit. One can see the roots of things he says now in things that he said at the very outset of his career in the early 1950s. But one might also speak, in the plural, of Chomsky's 'philosophies'. His thought has never been static, and within this unity there have been many important shifts of emphasis, many innovations and much reshaping of old theory into new. On some central issues, notably on the place of semantics in grammar, his views have changed not once but twice. For a historian of linguistic theory it is fascinating to trace the continuities and discontinuities in Chomsky's ideas. But for a student of current theory it is the best and possibly the only way to understand him. He is not a systematiser, and attempts to impose a system on him are liable to be betrayed by the next book that he writes. For those who are maddened by such things, he can be maddeningly inconsistent. At present, as always, his theories are in transition. To appreciate why they are going where they are one must have a thorough critical appreciation of their background.

I have also referred to Chomsky in particular, and not, in general, to a Chomskyan school. For it is doubtful whether any permanent school can be identified. Since the early 1960s Chomsky has, at any time, had crowds of followers. Many pupils have clung to his coat tails and, after publishing a thesis which was proclaimed to be important, have done little or nothing thereafter. Others have been scholars of independent intellect whose work has then diverged so much from Chomsky's own that no community of interest has remained. The greatest number have been teachers; by the early 1970s there were classroom versions of what Chomsky and others were supposed to have established in the 1960s which, as the decade wore on, were increasingly enshrined in textbooks. But both teachers and textbooks were left stranded when it was clear that he had taken a fresh turn. In the 1980s there is a new wave of followers, and little dialogue between them and the best of the old. We will refer to some of these people as we go along. But in Chomskyan linguistics the only element of continuity is Chomsky himself.

His career may be divided into four periods. Externally it is one: he moved as a young man from the University of Pennsylvania, via Harvard, to the Massachusetts Institute of Technology, and has stayed there since. But the first stage of his intellectual history begins in the early 1950s and is centred on his first book, *Syntactic Structures* (1957). In this period he was still strongly influenced by older theorists in the United States, retaining many of their biases while, in other ways, reacting against them. The second period begins towards the middle 1960s. It was brief, but immensely productive: a space of three years saw two monographs on grammar (1965a, 1966a), a rash excursion into the history of linguistics (1966b), an important set of general lectures (1968), not to mention a joint work on phonology (Chomsky and Halle 1968).

For many commentators this is Chomsky's classic period, the period of what he himself has called the 'standard' theory of transformational grammar. But by the end of the 1960s we can already distinguish the beginnings of a period of groping and reorientation, which was to last through most of the 1970s. This is marked most clearly by a series of technical papers (collected in Chomsky 1972a and 1977a) and a further set of general lectures (1976). By the end of the decade the reorientation was complete, and we may therefore distinguish a fourth phase whose latest manifesto (1986) opens, in part, a new perspective.

I will take these periods in turn. But this is not a chronicle, and I will not hesitate to refer both backwards and forwards where connections can be drawn.

1. 'SYNTACTIC STRUCTURES'

One remark of Chomsky's that seemed provocative or puzzling at the end of the 1970s was his assertion that the notion of a grammar is more central than that of a language (1980: 126ff). Since then he has changed his terms: what was formerly a 'grammar', and had been called so for the previous twenty years, is renamed a 'language' or 'I-language' (1986: 21ff.). But, in ordinary parlance, a grammar is not a language. It is merely one of the means by which a language, as the primary object of study, is described. Nor would Chomsky have disagreed with this at the beginning. To understand why both his views and his terms have shifted, we have to go back to his first book, and in particular to ideas that he inherited from his teachers.

In the view that was dominant in America when he entered the subject, the first or only task of linguistics was to study the formal patterning of units. For example, there is a unit *hat* which is identified by the smaller units /h/, /a/ and /t/, in that order. Ignore its meaning; in this conception of linguistics it is not relevant. There is also a unit *coat* and, still ignoring meaning, these can generally be substituted one for the other: *I wear a hat/coat, Some hats/coats were on sale*, and so on. In the key term of this school, *hat* and *coat* have similar DIS-TRIBUTIONS. We can therefore class them together, and can then go on to class together larger units such as *a hat* or *a coat, these coats* or *that scarf*, still for no other reason than that, in such sentences as *A hat would look nice* or *These coats would look nice*, they can all be said – meaning once more apart – in the same context. The description of a language is complete when the distribution for all units has been stated in terms of classes which are ideally general.

This approach was developed most consistently by Zellig Harris, in a book (1951) with whose typescript Chomsky himself helped (preface, v). Chomsky said later that this was how he learned linguistics (reference in Newmeyer 1980: 33). His own work shows this very clearly. Critics of Harris and others

had asked how a language could be described without appeal to meaning; but in Chomsky's view the implication that it could be done '*with* appeal to meaning' was 'totally unsupported' (1957: 93). He saw 'little evidence that "intuition about meaning" is at all useful in the actual investigation of linguistic form' (94). His own investigation of syntax was 'completely formal and non-semantic' (93), and linguistic theory in general, for him as for Harris, was a theory of distributional relations.

For Harris, a language was simply the collection of utterances whose formal structure one set out to describe. Similarly, for Chomsky, it was 'a set . . . of sentences' (1957: 13). In describing a language one must then do two things. Firstly, one must define the membership of this set. For example, the set 'English' has among its members *I wear a coat, That scarf would look nice*, and so on. In Chomsky's terms, these are GRAMMATICAL SEQUENCES of elements, whereas **Coat Wear I a* or **Would nice look that scarf* are sequences that are UNGRAMMATICAL. Secondly, one has to indicate the structure that each sentence has. For example, in *I wear a coat* the pronoun *I*, classed by Chomsky as a Noun Phrase, is followed by a Verb and a further Noun Phrase, which in turn consists of an Article plus a Noun. According to Chomsky, a *grammar* was a 'device' which performed both tasks. It contained a series of rules for the distribution of smaller and larger units. Thus, by one rule, a Noun Phrase can consist of an Article followed by a Noun. Unless there are other rules to the contrary, this excludes the possibility that successive Articles and Nouns might not form a Noun Phrase, or that, within such a phrase, the Noun might come first and the Article after it.

In this conception, a language is the primary object and a grammar is a set of statements about it. One standard way of putting this was to say that grammars were *theories* of languages. But let us now ask what it means to 'know a language'. As Chomsky saw it, speakers of English know what sequences of elements are grammatical sentences in English. But that is because they know the rules by which sentences are formed; to use a term which Chomsky popularised in the 1960s, it is because they have INTERNALISED (1965a: 8) the grammar of English. 'Knowing a grammar' is thus the primary concept, and 'knowing a language', in the technical and rather unnatural definition of a language with which he began, is at best derivative. It took several years for the implications of this shift to sink in. But once it had, it was obvious that this definition of a language made sense only when linguistics was restricted to the study of distributional relations. For these may indeed be seen as relations in a set of sentences. To 'study language' in a real sense is to study something else; and that might very appropriately be called an INTER-NALISED LANGUAGE or 'I-LANGUAGE'.

In the rest of this section we will look further at Chomsky's thought as we find it in his first phase. As we have seen, he followed Harris in excluding

meaning from the analysis of a language. The reason he gave was that there is no one-to-one relation between meaning and form. Forms can differ phonemically but mean the same; equally the same form can have different meanings. Not all morphemes have an independent meaning, and some forms that are not morphemes do. There is no coincidence between syntactic constructions such as Verb plus Object and constructional meanings such as Action-Goal (1957: 94ff.). Therefore a grammar had to treat forms on their own.

If a grammar was a theory of a particular language, a linguistic theory was in turn a general theory about grammars. But what can we expect of such a theory? The answer, in part, was that it had to specify the forms that grammars might take. They consisted of rules: thus, in Chomsky's formulation at that time, of phrase structure rules followed by transformational rules followed by morphophonemic rules. These rules were seen as *generating* the sentences of a language, in that, by following them through, it would be possible to produce any grammatical sequence of elements and none that were ungrammatical. Such rules had to be precise and had to conform to a format which the theory of grammar laid down. They also had to be as restrictive as possible. The main thrust of Chomsky's work in the 1950s was to demonstrate that some forms of grammar were too restrictive. With a finite state grammar (1957: Ch. 3) one could not generate the sentences of English. With a phrase structure grammar one might be able to generate them, but one could not describe their structure satisfactorily. With a transformational grammar one could do both. But one did not want to form a grammar which would also allow one to generate sets of sentences which were quite unlike any human language. Part of Chomsky's insight was to see this as a problem of mathematical formalisation. A grammar was a type of mathematical system. If the sets of sentences that can be generated by one type of system (A) include all those that can be generated by another type of system (B) but not vice versa, A is more POWERFUL than B. What was needed was a theory that had just the power – no more, no less – that was needed.

But a linguistic theory also had to provide what Chomsky called an EVALUATION MEASURE. Suppose that we have two grammars, both in the form that the theory prescribes and both generating the same language. But one may be simpler and, in that respect, better. According to Chomsky, the theory itself should then discriminate between them. Given a precise account of the forms of rule that it permits, including a detailed specification of the notation in which they are to be written, it should, in addition, prescribe a way of measuring the relative simplicity of alternative grammars for the same set of sentences. Now since these grammars are different they will in part describe the language differently. They might establish different units: for example, in *A hat would look nice*, one grammar might relate *would* to a Complement *look nice* while the other might say that *nice* was the Complement

of a single Verb *would look*. If not, they would establish different classes. For example, one might class both *I* and *a hat* as Noun Phrases, while the other might deal with Pronouns separately. The evaluation measure will therefore tell us which analysis of the language a given theory favours.

This account of the aims of linguistic theory was new and brilliant. But, in retrospect, it seems clear that there were problems. Grammars, as we have seen, were theories of languages and, like many other theories, they were based on limited evidence. They therefore made predictions: in Chomsky's words, which echo those of Harris (1951: 13) or Hockett (1948), any grammar 'will *project* the finite and somewhat accidental corpus of observed utterances to a set (presumably infinite) of grammatical utterances' (1957: 15). It was then true to the extent that its predictions of what was and what was not grammatical were borne out. But then we have on top of that another theory which will take two grammars that are in this sense equally true, and literally calculate that one is, in some other sense, better. Does 'better' just mean 'simpler'? That is what Chomsky seemed to be saying, and still more his associate Morris Halle (1961). But simplicity is not itself a simple notion: how then could we decide what sort of simplicity should be measured? Or does 'better' again mean 'truer'? Then in what respect truer and why should these levels of truth be separated?

These questions were answered, as we will see, in Chomsky's next phase (see section 2). For the moment, however, a more obvious problem was whether the study of forms and meanings could sensibly be divorced. For although Harris and others had sought to base their analyses on purely distributional evidence, they did not, of course, maintain that meanings could not be investigated. Likewise, for Chomsky, 'the fact that correspondences between formal and semantic features exist . . . cannot be ignored' (1957: 102). All that was claimed was that the formal features had to be investigated first, that descriptive linguistics (Harris 1951: 5) was concerned with them alone, and that any study of meaning had to come later.

In Harris's terms, the meaning of utterances was, 'in the last analysis', their 'correlation . . . with the social situation in which they occur' (1951: 187). This had its roots in Leonard Bloomfield's theory (1933: Ch. 9). For Chomsky, 'the real question that should be asked' was: 'How are the . . . devices available in a given language put to work in the actual use of this language?' (1957: 93). A language could therefore be studied like an 'instrument or tool' (103). On the one hand, we can describe its formal devices without reference to their use. In the same way, to develop the analogy, one could in principle describe a knife – handle, blade, sharp edge and all – without knowing, or while pretending that one did not know, that it was used for cutting. However, these devices have a purpose. So, given this account of the handle, edge and so on, one could then go on to incorporate it in a wider form of description which would

also explain what they are for. In the same way, we can envisage a 'more general theory of language' (102) of which a linguistic theory, in the sense already described, is only one part. The other part would be a separate 'theory of the use of language'.

In this light, both a grammar and a linguistic theory can be evaluated on two levels. Considered on its own, grammar A may be simpler than grammar B. This notion of simplicity may be given a precise sense, as we have seen, by an evaluation measure. In a looser sense, theory A may also permit a simpler form of grammar than theory B. Thus, in his first book, Chomsky argued that a theory which included transformational rules allowed a simpler grammar of English than one which included phrase structure and morphophonemic rules alone (1957: Chs. 5 and 7). But if we then go on to study meaning, simplicity is only one criterion. For we can also require of a grammar that it should 'provide explanations' (1957: 85) for semantic facts. The form /əneym/ has two meanings ('a name' and 'an aim'); this is explained, as Chomsky saw it, by a formal grammar in which it is divided into two different sequences of morphemes. In a passage that became famous, he argued that *the shooting of the hunters* could be used either of hunters shooting or of hunters being shot. That could be explained by a grammar in which, for reasons of pure simplicity, it is derived by different transformations (88f.). A theory which allows transformations is therefore better for another reason. Not only does it give a simpler description of the knife; but, if we may continue the analogy, a description which is simpler in terms of the proposed evaluation measure will also explain why the knife is held as it is and used to cut things.

What then was the real argument for transformations? For most of Chomsky's followers, it was precisely that they threw light on distinctions and similarities of meaning. On the one hand, forms which are ambiguous would have analyses to match, thus *the shooting of the hunters* and many other stock examples. On the other hand, a transformation could relate forms that were partly or wholly synonymous. For example, Actives were said to be synonyms or paraphrases of the corresponding Passives. Moreover, given that a linguistic theory allowed transformations, how did one decide in particular cases whether such a rule should be established? The sophisticated answer was that this should be decided by the evaluation measure; and, since the linguistic theory of which the measure was part could itself be seen as part of a more general theory which would also include a theory of use, it should ideally be so devised that a grammar whose formal descriptions contributed to the explanation of meanings would be simpler than one which did not. But in practice most of those who applied the model took what in an earlier phase of distributional linguistics might well have been disparaged as a 'short cut'. If there were semantic reasons for establishing a transformation they established it. The grammar might in an intuitive sense be simplified or it might

not; but the appeal to meaning was overriding.

Now Chomsky's followers are not Chomsky himself, and by the end of the 1960s this had led to a remodelling of grammar under the name of generative semantics (see the beginning of section 3) which he rejected. But neither he nor anyone else made any serious attempt to justify a syntactic evaluation measure. A proposal was developed in morphophonemics or, as it was misnamed, generative phonology. But in that field meanings were irrelevant and, even then, it did not in the end work. In syntax, despite the great place that it had in Chomsky's initial programme, the evaluation measure was still-born. For, by relating theories of form to subsequent theories of meaning, he had ensured that it would be transcended.

2. THE 'CLASSIC' CHOMSKY

In his account of the battle of El Alamein, Liddell Hart (1970: 315) comments on Montgomery's 'adaptability and versatility' in devising a fresh plan when his initial thrust had failed. It was 'a better tribute to his generalship' than his own habit of talking as if everything had gone as he intended. One might say much the same about Chomsky, both in his next phase and in the long re-adjustment which followed. From his own accounts, one might suppose that his thought has been consistent from the beginning. But in this way his true genius has often been disguised from his own troops.

Of the changes that mark Chomsky's general thinking in the middle 1960s, the most straightforward, on the face of it, was his extension of the concept of a grammar to include a SEMANTIC COMPONENT. Its syntactic component, as before, said nothing about meanings. Syntactic rules continued to indicate which sequences of morphemes could and could not represent grammatical sentences. But each sentence was now interpreted semantically. A generative grammar, as Chomsky put it in a series of lectures delivered in the summer of 1964, became 'a system of rules that relate [phonetic] signals to semantic interpretations of these signals' (1966a: 12). The objects that it generated were sentences in the old sense. But they now had meanings attached. More precisely, therefore, they were pairings of a phonetic representation of a sentence and its SEMANTIC REPRESENTATION.

How does this relate to the earlier division between a theory of form and a theory of use? One might say simply that the term 'linguistic theory' had been redefined: whereas it was previously part of a 'more general theory of language' (Chomsky 1957: 102), it now *was* that theory. But then there is a problem as to what was meant by 'use'. In 1957 Chomsky had talked of the 'actual use' of the language; this could be taken to mean that semantic theory was concerned with the use made of a particular utterance, by a particular speaker, at a particular time, in a particular set of circumstances. But a

generative grammar is a system of rules; particular uses vary indefinitely; there-fore, if a grammar was to assign semantic interpretations to sentences, these had to be something else. In Chomsky's formulation, they were 'intrinsic meanings' of sentences (1968 = 1972b: 71). In this context he no longer spoke of 'uses'. But, if we go back to the analogy of the knife, we might say that its intrinsic use is for cutting. I may then use it, on a particular occasion, to slice this particular cabbage which is in my kitchen. On another occasion I may use it in a non-intrinsic way, say as a makeshift screwdriver. In the same sense there was now a distinction that had not existed previously in Chomsky's thinking, between the meaning of a sentence as defined by rules and its actual meaning in a concrete utterance.

With that insight in mind, we can now turn to his general concept of 'knowing a language'. In his earliest publications Chomsky had said little about the psychological status of rules, his primary aim being to account for distributions. But at least one commentator had gone further. Towards the end of an enthusiastic review of Chomsky's first book, Robert Lees talked of the 'device' within the speaker's head which is used 'to generate the sentences of his language'. We cannot study it directly; but if our rules are adequate and general, then by the canons of science as Lees conceived them 'it is not too much to assume that human beings talk in the same way that our grammar "talks"' (Lees 1957: 406 ff.). In his own chapter on finite state grammars, Chomsky remarked that, if we accept that form of grammar, 'we can view the speaker as being essentially a machine of the type considered'. 'In producing a sentence', he too 'begins in the initial state, produces the first word of the sentence, thereby switching into a second state', and so on (1957: 20). Now Chomsky did not talk similarly about machines which included phrase struc-ture rules and transformational rules. But to Lees at least it seemed that a grammar was a literal model for the production of utterances.

Two years later Chomsky dismissed the suggestion (1959: 56). But at the same time he assumed, without argument, that a generative grammar could be said to 'characterise abstractly' what he later called the speaker's linguistic COMPETENCE. Speakers can, for example, 'distinguish sentences from non-sentences'; as Chomsky saw it, that ability is characterised by a grammar that gives rules for the distinction. It also characterises, 'in part', their ability to understand a sentence that they have not heard before. In his words they are 'somehow capable of determining the process by which this sentence is derived' in the grammar. Likewise it can characterise their ability to 'note certain ambiguities'. Now a language, as we have seen, was a set of sentences; and a speaker who knows the language can be said to know what these sentences are and to know their structure. Accordingly, he can be said to know a grammar: that is, he knows a set of rules which specify what the language is. In this light, Chomsky uses the term 'grammar' with what he later

called a 'systematic ambiguity' (1965a: 25). 'Grammar₁', we might say, is a set of rules constructed by a grammarian. But in Chomsky's interpretation these are an attempt to characterise the competence of a speaker, and that is itself a grammar. 'Grammar₂' is thus the set of rules, that everyone who knows a language has also 'in some sense constructed' (1959: 57).

If linguistic competence is 'the speaker-hearer's knowledge of his language' (1965a: 4), 'the actual use of language in concrete situations' constitutes his PERFORMANCE. A generative grammar cannot account for this directly: thus, to return to semantics, it can account for intrinsic meanings but not actual, concrete meanings. Nor was it seen any longer as a projection from a set of 'observed utterances'. On the one hand, Chomsky remarked that 'a record of natural speech will show numerous false starts, deviations from rules, changes of plan in mid-course, and so on' (1965a: 4). A grammar was not concerned with these, but only with an ideal form of speech in which all sentences were correct. On the other hand, he proposed that certain sentences which were grammatical might, in performance, be unacceptable (10ff.). Again the grammar was concerned with grammaticality only. Never-theless a speaker's performance rested on his underlying competence, and it was this competence that a grammar (grammar₁) described. Therefore, in any study of performance, the study of a grammar (grammar₂) had to be primary. As Chomsky had put it in his earliest formulation, any 'direct attempt to account for the actual behavior' of speakers or hearers, 'not based on a prior understanding of the structure of grammar', will have 'very limited success' (1959: 58).

All this was quite a mouthful, and it is remarkable, in retrospect, that Chomsky should have introduced it with so little argument. He seems genuinely to have believed that, if one was prepared to think about the psychology of language at all, what he had said was uncontroversial. But once a 'grammar₁' is reified as a 'grammar₂', the rest of Chomsky's mature theory follows without much difficulty. A speaker has as a child acquired, constructed or internalised grammar₂; to be able to do so, children must have in their heads a LANGUAGE-ACQUISITION DEVICE which takes 'primary linguistic data' (1965a: 25, 31) as input and yields a grammar as output. There is therefore a direct comparison, developing in effect a remark of Hockett's (1948), between the construction of a grammar₂ by a child and that of a gram-mar₁ by a linguist. In either case, grammar₁ or grammar₂, the set of rules is very complex. Moreover, in the child's case, its construction is 'accomplished in an astonishingly short time, to a large extent independently of intelligence, and in a comparable way by all children'. (See Chapter 10, below, section 4.) How can these 'facts' (1959: 58) be explained?

Chomsky's answer was to reify not just the concept of a grammar, but also that of a linguistic theory. As first envisaged, this was a second-order theory

that restricted the forms that grammars might take and the class of languages that they might generate. A grammar was, in turn, a first-order theory about a language. But let us now suppose that such restrictions are known to children when they learn their native language. In that case, just as a grammar$_1$ is an attempt to describe a grammar$_2$, so a linguistic theory posited by a linguist – call it linguistic theory$_1$ – can be reinterpreted as a hypothesis about a linguistic theory$_2$ (Chomsky again makes clear that he is using terms with 'systematic ambiguity') that every child possesses. All children must possess it equally. It constitutes a faculty of the mind distinct from general intelligence, and therefore stupid children can acquire a grammar as quickly and successfully as bright children. By the same token, it cannot itself be learned. Instead it must be part of our genetic make-up; briefly, it must be *innate*.

The linguistic theory which was reified in this way was conceived in other respects exactly as in 1957. First, it specified the form that grammars might take. As Chomsky saw it, 'a child who is capable of language learning must have', among other things, 'some initial delimitation of a class of possible hypotheses about language structure' (1965a: 30). This was again seen as restricting the class of languages for which a grammar might be constructed. Accordingly, 'the child approaches the data with the presumption that they are drawn from a language of a certain antecedently well-defined type', and must then 'determine which of the . . . possible languages is that of the community in which he is placed' (27). In short, he already knows what human languages are like and what grammatical rules are like. 'Language learning would be impossible', Chomsky said, 'unless this were the case' (27). The passages cited make clear how complete the parallel was thought to be, not just between theory$_1$ and theory$_2$ or grammar$_1$ and grammar$_2$, but in the entire cognitive task that children and grammarians faced.

Second, the theory had to provide an evaluation measure. In Chomsky's words again, 'a child who is capable of language learning' must, in addition, have 'a method for selecting one of the (presumably, infinitely many) hypotheses that are allowed . . . and are compatible with the primary linguistic data' (30). In this light he was able to explain more clearly what a linguist's evaluation measure – evaluation measure$_1$ – was meant to assess. A grammar$_1$ is, once more, a hypothesis about a speaker-hearer's grammar$_2$. It is therefore *descriptively adequate* (1965a: 24) 'to the extent that it correctly describes' the competence that underlies his performance. But in developing their competence, or in constructing a grammar$_2$, children have 'a method of selecting' between hypotheses. Accordingly, the linguist's evaluation measure$_1$ may be interpreted as a theory about this method. In devising it, we again aim to select the simplest and most general set of rules; however, we do this not because we have an *a priori* concept of elegance, but because we assume that there is an innate evaluation measure – evaluation measure$_2$ – which selects a

grammar in the same way. 'Simplicity' was therefore reinterpreted as an empirical concept (see, in particular, Chomsky and Halle 1965). The evaluation measure literally measured truth – that is, the descriptive adequacy of grammars$_1$.

The chapter in which these ideas are introduced is very loosely argued (see Matthews 1967: 121 ff.). The ideas have always been hard to expound, and I have therefore given quotations where possible. But it was clear at the time that their historical importance was much greater than the trains of thought by which they had been reached. For once we accept that a linguist's grammar$_1$ is a description of a speaker-hearer's grammar$_2$, and a linguist's linguistic theory$_1$ an account of a learner's innate linguistic theory$_2$, our subject is given a new purpose and a new standing in relation to other disciplines. In later years, Chomsky and his followers were to talk resoundingly of language as a window on the mind, of a linguistic faculty peculiar to human beings and unparalleled in other species, of the problems raised for human evolution, of linguistics in general as a science whose findings no other human science, from philosophy to biology, could ignore. But although these external prospects seemed to many to be very exciting, perhaps the most important implications bore directly on the discipline itself and the methods by which its findings could be reached.

Let us consider first the data on which our descriptions of languages are based. In Chomsky's earliest phase, a grammar (grammar$_1$) characterised a set of sentences. The data were therefore possible sentences, either observed utterances or sentences that could in principle be utterances. They formed a corpus or sample of the language, and the grammarian's task was to extrapolate from the part to the whole. Where there were alternative extrapolations he chose the simplest. If he was not sure whether a sentence was possible or not, he started from data that were certain and made whatever extrapolation offered simpler or else more general rules. In Chomsky's own words (1957: 14), he was 'prepared to let the grammar itself decide'.

But now a grammar$_1$ is a description of the speaker-hearer's competence (grammar$_2$). As such, it accounts for a variety of abilities: thus, as we have seen, the ability to 'distinguish sentences from non-sentences', to 'note certain ambiguities', and so on. Moreover, a person's competence is reflected only indirectly in performance. Suppose, for example, that a speaker of English is observed to say *I went to home*. It may be that the observation is misleading and that he really meant to say *I went home*. Or perhaps he was going to say *I went to the pub* but changed his plan too late. Or perhaps he really did say *I went to home*, but the preposition was used by mistake. The observation itself is unreliable, and a corpus which includes it may well be an inaccurate sample of the language whose grammar$_2$ has been internalised. We will therefore do better if we simply ask the speaker, or attempt to find out by

some other direct experiment, whether *I went to home* is grammatical for him. Since he has internalised a grammar₂ he knows whether it is or not, and it is this knowledge, not his actual speech, that we are seeking to describe.

Now consider a sentence like *I watched the shooting of the hunters*. A speaker who has internalised a grammar₂ of English knows that this is ambiguous. He knows that it has two different semantic representations and, corresponding to these, two different syntactic structures. There is therefore no need to argue, as before, from distributional evidence. Nor is there any reason to appeal, in either the first or the second instance, to data bearing on the actual use of such a sentence. Nor does it make sense to ask whether our evidence is of form or of meaning. It is simply evidence of the speaker-hearer's internalised knowledge of his grammar₂, and, in constructing our own grammar₁, we may rely directly on it.

In short, a grammarian's data are primarily the speaker's *intuitions*. That they were among his data was not new: Lees's review of Chomsky's first book is again more explicit than the book itself (Lees 1957: 376). But increasingly they became the only evidence that generative grammarians were to use. Nor were they got from what had earlier been called 'naïve informants'. A follower of Chomsky typically worked on his own language; he himself spoke it, and therefore had his own intuitions about what was grammatical and ungrammatical, which sentences were ambiguous, and so on. Therefore he could proceed by pure introspection, without appealing to observational evidence at all. By the end of the 1960s this method was employed with almost total confidence. When one grammarian's intuition clashed with that of another, either each said he was right and the other was wrong, or they agreed charitably that the 'dialects' which they had learned as children must be different. Scholars began to worry about the problems of investigating dead languages, for which such data could not be got. Others argued that only native speakers could describe a language safely, since only they had intuitions which were correct. In his review of 1957, Lees had distinguished carefully between this form of evidence and 'the intuitive or prescientific perceptions which the linguist, qua scientist, has about the data'. But, in practice, it became very hard to keep them apart.

A second important bearing was on the study of UNIVERSAL GRAMMAR. It had always been assumed that there were features common to all languages; we could therefore 'look forward', as Bloomfield (1934 = 1970: 285) had put it, 'to . . . a General Grammar, which will register similarities' between them. The term 'linguistic universal' or 'universal of language' was itself introduced by scholars who were not among Chomsky's followers (see especially Greenberg 1963; and, here, Chapter 9 below). But Chomsky's new interpretation of linguistic theory quite transformed this field of research.

A child, to recapitulate, constructs a grammar. This is very complex; so, to

construct it as consistently and quickly as they do, children must already know in detail many features that a grammar has to contain. But they are not genetically equipped to learn particular languages. A baby born of parents who speak English can as readily internalise a grammar of Russian or Chinese or Quechua if that is spoken in the community in which it grows up. Therefore the features that are known innately must be common to all languages. At the same time they must be specific. If they were merely general indications – as, for example, that a sentence can be analysed into words, or that it has a deep structure and a surface structure – their value as a blueprint for constructing grammars would be slight. Therefore languages must be much more similar than Chomsky's predecessors had supposed. The diversity of their structures must be superficial. Behind it a rich and intricate set of universal principles must be waiting to be discovered.

This was a matter of faith and not an empirical finding. It simply followed from the logic of Chomsky's new conception of linguistic theory. Let us now consider how a universal feature may be discovered. The obvious method, if one was not a follower of Chomsky, was to look in general at the widest possible range of languages. This requires extensive knowledge and, to Bloomfield at least, it had seemed that 'lack of data' still forbade it. But in Chomsky's programme this form of study was in any case peripheral. For one cannot argue directly from universality to innateness. As critics pointed out, a feature may in fact be present in every language but may not be genetically inherited (thus, for instance, Matthews 1967: 122 ff.). But, conversely, if a feature is innate it has to be innate universally. To discover it we do not have to look all round the world. We must simply show that adult speakers of whatever language could not construct the grammars that, on the evidence of their abilities, they do construct unless, when they are learners, this feature is already fixed. Indeed the evidence of just one language may suffice. 'Paradoxical as it may seem at first glance', Chomsky remarks in a note to the next chapter (1965a: 209), 'considerations internal to a single language may provide significant support' not just for its own grammar but also for universal theory.

I recall that at least one reader flung the book down when he read the note which I have cited. But in retrospect its wording was cautious. For, in the years that followed, and for reasons that were perfectly legitimate if one accepted the logic of Chomsky's theory, the study of the universal properties of human language was to proceed almost wholly on the basis of a generative grammarian's intuitions about English.

3. THE PERIOD OF TRANSITION

Many years later, Chomsky said that if he had to rewrite the introductory chapter of *Aspects of the Theory of Syntax*, he would not change what he had

written. It is indeed the pivot on which the 'generative enterprise' (Chomsky, Huybregts and van Riemsdijk 1982) was set in motion. But although the core of its interpretation of grammars remains, there are other things in it, perhaps less central to Chomsky's own evolving concerns, that he soon abandoned or qualified.

Let us begin once more with meaning. The semantic theory which Chomsky took as uncontroversial in the mid 1960s was not his invention. It was primarily that of Jerrold Katz (Katz and Fodor 1963, Katz and Postal 1964), and its central tenets, as we have seen, were that sentences have what Chomsky called intrinsic meanings, and that these are derived by rules interpreting syntactic structures. A more precise way of putting this was to say that each syntactic structure 'uniquely determines' (Chomsky 1966a: 13) a semantic representation. But such a theory must provide criteria for distinguishing intrinsic meaning from all other meanings that a sentence may have when it is uttered. Take once more our analogy of the knife. A knife that I have often used for chopping onions was used by my mother as a bread knife. Has it just one general use (cutting)? But surely there is some intrinsic difference between this knife, which is large, and one which I might use to peel an apple. Has it, alternatively, two intrinsic uses (chopping, slicing)? But how does one put a stop to the distinctions that might then be drawn? Thus peeling an apple is not the same as slicing a cucumber; yet I often do both with the same knife, which is still intrinsically unsuitable for slicing bread. Do we say that the large knife has intrinsically just one specific use (*either* chopping onions *or* slicing bread)? In that case either my mother has used it, or I have used it, in a deviant way. But which of us?

A theory of intrinsic meanings cannot avoid similar problems, and in the early 1970s they were beginning to cause theoretical anguish. One reason is that most of Chomsky's followers had adopted the theory of GENERATIVE SEMANTICS. According to this, a generative grammar began by characterising semantic representations. These were not determined by, or derived by the interpretation of, syntactic structures. Instead the latter were derived from them. Therefore, in dealing with any body of data, a grammarian's first task was to work out what the semantic representations should be. Was a sentence ambiguous, and if so how many meanings did it have? Answers were given which seemed at first sight to be ludicrous: for example (seriously), that any plural had an infinite set of semantic representations (*men* = 'two men', 'three men', 'four men' and so on); or, in parody, that *He stood on one leg* was ambiguous because it could have been his left leg or his right. Ludicrous such proposals may have been. But they rest on judgements that, at some level, we can recognise to be correct, and if they do not concern intrinsic meanings, one is forced to wonder how such meanings can be teased out.

Chomsky himself dismissed generative semantics. At the time many followers were surprised. For he himself had said that semantic representations were determined by rules; in principle, there was no reason why they should not be generated directly. He had also said that judgements of ambiguity and so on were data that reflected a speaker's competence; and, as we noted in section 1, semantic arguments had been increasingly used as a primary ground for establishing transformations. The more that was done, the more grammarians were led directly to meanings and not merely to a deeper level of syntax. Finally, in the lectures of 1964 in which he had introduced the revision of his model of grammar, Chomsky had implied that the reason for starting from syntax was simply that we did not know very much about how meanings should be represented (1966a). That was a clear challenge to find out more and start from them instead.

But Chomsky had originally taken a different view of the relation between grammar and meaning and, although he had approved the concept of semantic representations, saying at one point that it was, 'Quite obvious that sentences have an intrinsic meaning determined by linguistic rule' (1965b = 1972b: 115), his own interests have always centred on syntax and he has rarely discussed a particular problem of meaning which did not have a syntactic point to it. Where other aspects are concerned (for example, the meanings of lexical units or the status of speech acts) he has said little and then mainly in polemic (as in Chomsky 1976: ch. 2). Moreover, there are signs that, by the end of the 1960s, he had himself begun to doubt that the intrinsic meanings of sentences could be isolated. In a lecture in 1969, he remarked that 'the notion "representation of meaning" or "semantic representation" is . . . highly controversial'. 'It is not at all clear', he goes on, 'that it is possible to distinguish sharply between the contribution of grammar to the determination of meaning, and the contribution of so-called "pragmatic considerations", questions of fact and belief and context of utterance' (1972b: 111). In short, it is not clear whether semantic competence can be distinguished from performance. At the time, he continued to posit semantic representations. But they were an abstraction, and might prove invalid.

Seven years later, in a series of conversations that did not appear in English until 1979, Chomsky effectively ditched Katz's theory. He points out, correctly, that it is not what he had proposed in the 1950s. In addition, it posits a semantic representation based on a 'universal system of semantic categories' (1979: 141). But although some 'traditional notions' can be taken as universal (for example, 'agent of action' or 'instrument'), and although some other features of meaning (for instance, anaphora or the properties of quantifiers) also belong to 'the system of rules that specifies our purely linguistic knowledge', it is 'not at all clear' that a universal system, which he himself had also taken as necessary (1965a, 1966a) could be defended. In the next paragraph he returns to the role of pragmatic factors. 'It is not at all clear', he says,

'that much will remain if we try to separate the purely linguistic components of what in informal usage or even in technical discussion we call "the meaning of linguistic expression"'. 'I doubt', he continues, 'that one can separate semantic representation from beliefs and knowledge about the world' (1979: 142).

This was a major turn-around, as Katz (1980) at once recognised. But where exactly does it lead? A first possibility, which no one at the time appears to have spelled out in so many words, would have been a return to something like the view that Chomsky had held in the 1950s. A generative grammar would be concerned with syntax and phonology, and the primary aim of syntax would be to describe distributions. It would also assign to sentences structures which were suitable for semantic interpretation. But the interpretation itself would lie outside the grammar, and in it many different factors, some depending entirely on the state of a particular speaker or hearer on a specific occasion, would be mingled.

But it is easy to see why that would not do. For in the conversation cited Chomsky accepts that some semantic notions (anaphora, roles of participants and so on) are universal. If so, they are candidates for the innately determined universal grammar (section 2) and, if they are part of that, they must be part of the particular grammars that speakers construct. Alternatively, they are candidates for some other innate mental structure that is also specific to language, and it is hard to see why this should be separate.

A second possibility would have been to abandon the concept of an internalised grammar. According to Chomsky, a grammar was, by definition, a set of rules relating meanings to phonetic signals (see again Chomsky 1966a: 12). But meanings are only partly determined by grammatical rules; accordingly, such a grammar is a contradiction. Moreover, it is 'not at all clear' that the contribution of rules can be separated from that of other factors. If it cannot, there is no other way in which a grammar with semantic rules may be delimited. If we follow this argument through, a speaker's competence is simply the ability to speak and understand speech in specific contexts. There would be no delimitable aspect that a grammar could be said to describe.

Some proponents of generative semantics had already reached this conclusion. But it plainly strikes at the foundation of Chomsky's philosophy. For it is because speakers were believed to have internalised a grammar (the grammar$_2$ of section 2) that a linguist's grammar (grammar$_1$) could be interpreted psychologically. And it is because the linguist's rules were so complex, and the speaker's were assumed to be similar, that the ability to learn languages had to be explained by a prior knowledge of universal grammar. And it is because the knowledge of grammar was believed to be separate from the use that speakers made of it that one could posit a specific faculty of the mind to which this prior knowledge belonged. Abandon the basic concept and all that rests on it dissolves.

Such arguments are a reconstruction, since Chomsky did not debate the matter overtly. But the alternative that he adopted was in effect a compromise. On the one hand, it posited that some semantic rules – call them semantic rules$_1$ – do apply independently of the contexts in which sentences are used. They include, for example, rules for obligatory anaphora (in *Bill cut himself* the reflexive pronoun must be anaphoric to *Bill*), for obligatory non-anaphora (in *Bill cut him* the simple pronoun cannot be anaphoric to *Bill*), for the meanings of agent, goal or instrument (in Chomsky's terminology these are 'thematic' relations), and for other features of grammatical meaning that Chomsky took to be universal. Such rules form what was at this point called a 'sentence grammar' (1976: 105). But, on the other hand, there is another type of semantic rule – call them semantic rules$_2$ – which operates conjointly with other forms of knowledge. An example is the interpretation of *him*, in *Bill cut him*, as referring not just to a male individual, but to that particular individual who, on a particular occasion, is in question. Semantic rules$_2$ are also part of our linguistic knowledge. But they form a second semantic component that, in a narrow sense, is outside the grammar.

Having adopted this theory, Chomsky was then free to concentrate on sentence grammar. Within it a first set of rules – the rules of syntax –.derived a structure that (still on the model of the 1960s) must be interpreted by semantic rules$_1$. For example, in *Who did he say Mary kissed?*, which is a type of sentence that Chomsky discussed throughout the 1970s, the syntactic structure shows, among other things, that *who* is moved by transformation from an initial position after *kissed*. The semantic rules$_1$ will then derive what Chomsky called the LOGICAL FORM of the sentence. This term was defined by the general theory: it referred to 'those aspects of semantic representation that are strictly determined by grammar, abstracted from other cognitive systems' (Chomsky 1977a: 5). But, as the name implies, a logical form particularly represented what older grammarians would have described as logical relations. Thus *Who did he say Mary kissed?* had the logical form 'for which person *x*, he said Mary kissed *x*?' (Chomsky 1976: 99). Apart from marking *who* as personal, this is in particular designed to show that it is logically the object of *kissed*.

But where was the division between a syntactic structure that determines logical relations and a semantic structure that represents them? The semantic rule that links *who* to its position in the subordinate clause is of a type appropriately called a RULE OF CONSTRUAL (Chomsky 1977a: 6). But do not rules of syntax also show how sentences are construed? They operate differently; but, in the same example, there is a transformational link between *who* in its initial position in deep structure (. . . *kissed who?*) and the same word in its position in surface structure (*Who . . . kissed?*). By what criterion is a construction in part syntactic and in part semantic?

Now syntax had originally been distributional (section 1), and in the model adopted in the 1960s (section 2) it had continued to distinguish sentences from non-sentences. But by the mid 1970s this constraint had been dropped. Take, for example, the non-sentence *They said that Mary kissed each other*. At the beginning of the decade, the grammatical *They kissed each other* was usually derived, by a transformation, from the deep structure of *They each kissed the other*. The transformation could apply within a clause; but, given the deep structure of *They each said that Mary kissed the other*, it could not cross the boundary of a clause to attach *each* to the object of the subordinate verb. That is still the solution assumed by Chomsky 1973 (= 1977a: 89ff.).

But another solution is to say that such a sequence cannot be interpreted. In *They kissed each other*, the reciprocal phrase is linked anaphorically to *they*: that would again be effected by a rule of construal. But this semantic rule may likewise be said not to apply across clause boundaries. So, in *They said that Mary kissed each other*, there cannot be an anaphoric link between *each other* and *they*. But the reciprocal phrase cannot be linked to *Mary* either, since one is plural and the other singular. Nor, finally can it be understood without an antecedent. It follows that the sequence is unconstruable; but then, if its unacceptability can be explained at that level, there is no reason why the syntax should not permit it.

That is the solution adopted in Chomsky's next paper (1975 = 1977a: 178). It is merely one of many cases (some already in Chomsky 1973) where a sequence once excluded by the rules of syntax is instead rejected because no logical form can be assigned to it. Chomsky accordingly denied that there was any *a priori* difference between levels. A speaker knows, for example, that *The police think who the FBI discovered that Bill shot* is ungrammatical, whereas, if we replace *think* with *know*, it is grammatical. But, as we remarked in passing in section 2, he cannot say directly whether this intuition is about form or about meaning. Chomsky himself was 'not persuaded that the question makes very much sense, or that any reasonably clear criteria exist to settle it' (1976: 95). The same type of fact might in principle be explained in either way.

At the time this matter did not seem to be central. For most commentators, the issue of the day was whether semantics was 'generative' or 'interpretive'. Should a grammar start from semantic representations (as in the model of generative semantics) or should they be derived from representations of syntax? If one took the 'interpretive' view, a second question concerned the level of syntactic structure that they interpreted. Originally it had been, by definition, the deep structure; subsequently, it was a paired deep structure and surface structure (Chomsky 1972a); later still, just the surface structure (Chomsky 1976 and thereafter). Successive models of grammar were distinguished on that basis: a 'standard theory' of the mid 1960s; an 'extended standard theory'; finally a 'revised extended standard theory'.

But in retrospect it seems clear that, behind the façade of technical progress, we were in fact witnessing the death throes of distributionalism. In Chomsky's earliest phase, the whole grammar was concerned with distributional relations only; and, as we saw in section 1, the requirement that it should be formally simple was prior to the expectation that it might, in part, explain the uses of sentences. But in the 1960s the priorities were reversed. A measure of simplicity, if relevant at all, applied to the entire grammar, and this included a semantic component. The primary requirement was that grammars should at all levels be descriptively adequate. Moreover, this did not imply that they should deal with every fact of distribution. For among the sentences generated by the grammar there might be many that, for other reasons, could not serve as utterances (thus again Chomsky 1965a: 10ff.). Their unacceptability would be explained by the interaction of a grammar, as a theory of competence, with a theory of performance.

Distributionalism died hard. In the enlarged grammar of the mid 1960s, the syntactic component was effectively equated with that of the earlier distributional model. It continued to characterise 'all and only the sentences' that were deemed to be grammatical, and, in justifying the form that its rules took, Chomsky appealed directly to earlier arguments. The semantic component was correspondingly no more than an interpretative appendage. As late as the mid 1970s, at least one textbook still insisted that formal arguments were separate from semantic arguments, and that the latter should not be used to justify syntactic rules (Culicover 1976: 45 and elsewhere). But by that stage Chomsky himself had concluded that the separation was nonsense. The very basis for a distributional grammar, or for a purely formal theory within 'a more general theory of language' (Chomsky 1957: 102), had collapsed.

Where did that leave the criterion of generativity? In the beginning the first requirement for a generative grammar was that it should generate 'all and only the sentences'. As Chomsky put it in the 1960s, it had to be *observationally adequate*. As the primary criterion for grammars this was superseded, as we have seen, by that of *descriptive adequacy*. But a still more vital requirement was that a theory of grammar should be *explanatorily adequate*: it should explain how a child's construction or development of a grammar is possible. Let us suppose then that a particular aspect of a speaker's competence has to be ascribed to a universal principle. We say 'has to' because we have evidence that speakers have internalised a certain set of rules, and cannot explain how they could do so if the principle were not innate. Now there is no objection if these rules as such are not observationally adequate. For no particular set of rules, and no particular component of the grammar, has a privileged role in separating what is grammatical from what is ungrammatical. Suppose, for instance, that the principle determines a set of transformations. It might allow numerous constructions that seem wholly ungrammatical. But perhaps they

are so because there are other rules and principles, perhaps unknown, which block the corresponding logical forms. Or perhaps they are excluded by the rules deriving phonetic forms. Or perhaps the explanation lies outside the grammar altogether. Just as we do not have to look at languages in general to propose that a feature is universal (end of section 2), so we do not have to be sure of every other aspect of the speaker's mind.

At the end of the 1970s Chomsky began to emphasise that the mind, as he saw it, had a *modular* structure. It should be seen not as an undifferentiated whole, but as a system of 'distinct though interacting' subsystems, each of which has its own properties and is 'organised along quite different principles' (1980: 40ff., 89). Our linguistic faculty had originally been conceived as one module; as such, it interacted with other modules, and it was only in that way that our actual use of language could be explained. But by the middle of the decade it too had a modular character. In the light of Chomsky's shifting view of meaning, it made sense, 'in particular, to distinguish what is sometimes called "grammatical competence" from "pragmatic competence"' (1980: 59). Within grammatical competence we can then conceive of further modules, distinguished not, as before, by *a priori* concepts of linguistic levels, but again by different organising principles. These too are 'distinct though interacting'. We cannot know in advance what they are, and we cannot expect of any one of them that it should characterise a language with observational accuracy.

The period of readjustment was then over. It had seen, in part, the rejection of ideas that had been innovations in the 1960s: in particular, Katz's notion of a semantic component. But it had also made clear the peripheral status of a set of notions that had been central in the 1950s. One was Chomsky's initial concept of a language: as we noted in section 1, it took time for the implication to work through, but once a grammar is reified as the speaker's competence or 'internalised language', the 'externalised language', as Chomsky now calls it, is 'an epiphenomenon at best' (Chomsky 1986: 25). In the 1960s Chomsky had reified the entire linguistic theory of his first phase. Thus, in addition, the sets of sentences called natural or human languages were of 'a certain antecedently well-defined type' (Chomsky 1965a: 27) and, in constructing rules that generated the one spoken in a particular community, a child made use of an evaluation procedure. One therefore had the illusion that the notation and the generative power of grammars were still important. But these too were relics of the earlier marriage between mathematics and descriptive linguistics. Fifteen years later they had in practice ceased to matter.

4. A NEW SYNTHESIS

The modular theory that has developed in the 1980s is essentially a theory of what Chomsky calls CORE LANGUAGE (or, in his earlier terminology, CORE GRAMMAR). This is one of a mass of burgeoning ideas that emerged obliquely in Chomsky's own work and which it is hard to separate and follow systematically. But, to put it briefly, a child has an innate universal grammar (section 2). This comprises a set of universal rules or principles, each of which allows an individual language – we are speaking again of the 'language' internalised by a speaker – to vary within limits. But let us now propose that, at a certain level of abstraction from the detailed facts of particular languages, the variation that the principles allow is finite. They will then constitute a set of PARAMETERS. Each parameter will have a fixed set of values; and, in developing his internalised language, a child will select, by experience, a particular value for each. The result is his core grammar (Chomsky 1981: 7) or core language (1986: 147): a central part of his knowledge that develops solely by a choice of values that are already innately given.

The term 'core grammar' had first been used in the later 1970s, when it referred in particular to an area of grammar delimited by certain specific principles (Chomsky 1977b). They included the principle by which, in a sequence like *They said that Mary kissed each other*, the rules of construal could not take *each other* as anaphoric to *they*. By 1980 this was part of what was called (temporarily) the 'opacity principle'. They also included a principle by which, for example, *Who did he believe the story that Bill kissed?* cannot be derived by transformation from an underlying *I believed the story that Bill kissed who*. From the early 1970s this had been known as the 'subjacency principle'. It thus united parts of what were technically semantic interpretation with other matters that were still conceived syntactically. In current versions, it includes 'such modules of grammar as X-bar theory, theta theory, binding theory, Case theory, control theory, and bounding theory' (Chomsky 1986: 155). Each 'theory' is highly abstract, and it is only by their interaction that the representations of a set of sentences within core language is determined. Within an internalised language, a core language is in turn no more than a fragment. There is therefore further interaction between this central set of subsystems and a PERIPHERY consisting of whatever else 'is added on in the system actually represented in the mind/brain of a speaker-hearer' (Chomsky 1986: 147). Finally there is a wider interaction, as before, between the systems that make up the speaker's internalised language, his pragmatic competence (see again Chomsky 1980), and other components of his mind. If we start from a common-sense notion of the speaker's knowledge of a language, core language is a very restricted and very abstract part.

The implications of this new approach were partly clear in 1981, when it

first crystallised, and are in part still emerging. But, to begin with the simplest, a core grammar has no rules. In acquiring a language, children have to learn the properties of individual words, including those properties that relate to universal grammar. The various parameters must also be fixed. But within the core language that is all: any specific rule, for any specific construction that is not allowed directly by the universal principles, must by definition fall outside it. As Chomsky makes clear, a core language is in this respect unlike the generative grammars that he had conceived of earlier. Universal grammar is, as it were, a system that is 'only partially "wired up"' (Chomsky 1986: 146). As soon as a child has fully wired it up, 'the system functions' and a core language is in being.

If there are no rules, there can be no types of rules. So, in particular, there is no distinction between rules that can be labelled syntactic and others that can be labelled semantic. Now the universal principles will still distinguish different levels of representation, and in recent work these are still named in ways that recall the model of syntax and semantics current ten years earlier. But since 1981 they have all been of the same sort. A sentence such as *Who did you see?* has an initial structure (roughly) *you saw who*, and this is naturally represented in the same way as the structure which results when *who* is moved (*who you saw*). Since the late 1970s these have been known respectively as the 'D-structure' and the 'S-structure'. A third level is that of 'LF', a term intended to 'suggest' (Chomsky 1986: 67) logical form. But here too there is no fundamental difference. The 'LF representation', as it is called, is another object of precisely the same kind as the D-structure and the S-structure. Indeed it may on occasion be identical to either (see, for instance, Chomsky 1986: 75 ff.) or, for that matter, to both. The form of representation that was originally called a logical form (see again Chomsky 1976: 99) is now called an interpretation of the LF representation, or an 'LF interpretation' (Chomsky 1986: 76 and *passim*).

The drift of all this is perhaps not perfectly clear. But LF is itself described as a syntactic level (e.g. Chomsky 1986: 84) and, on the face of it, the whole of the core language is concerned with syntax in a traditional sense. It is not, of course, a distributional syntax, or the relic of distributional syntax that had survived in the 1960s. But a theory of core grammar allows certain sets of constructions, all of which are represented by an LF representation, a D-structure and an S-structure. The choice of parameters will select a particular set for any particular language. At the same time, individual words will have specific properties. For example, *each other* is a reciprocal pronoun: therefore it can only satisfy constructions where the universal principles allow it to be linked to a plural antecedent. Properties like 'reciprocal' are traditionally semantic, and Chomsky too describes the lexicon in terms of 'semantic selection' (1986: 86). Beyond this, and beyond the internalised language as a whole,

semantics can be seen as a relation between language and the world (44) or between language and other cognitive systems (68). But within the language semantic construal is essential to syntax and not separate from it. That was the view before distributionalism, and for Chomsky too it now seems that it may be so.

Another implication, which was clear much earlier and has a far more central place in Chomsky's programme, concerns the extension of the theory of universal grammar to languages other than English. As we noted in section 2, the study of a single system can be instructive. In the middle 1970s Chomsky reaffirmed this point (1976: 118), and five years later saw its denial as 'irrational' (1981: 6). But by this stage proposals were becoming complex, and their limitation to English, which was virtually complete until the brink of the 1980s, was a 'serious limitation' (1976: 118). For suppose that we have developed a theory that, in the case of English, has a wide explanatory power. We then find that there are other languages for which a different theory is needed. Perhaps it will have partly similar principles or perhaps ones that are different altogether. We do not want to say that either theory is false, since our original data, which by the logic of the argument required us to posit that these structures were innate, are left unexplained. But if they are innate they must be innate universally. How then can both theories be true?

The concept of parameters provided an immediate answer. Suppose that principle A, which holds for one set of languages, differs in only one respect from principle B, which holds for another. We can then say that A and B are the same principle; but it incorporates a variable with two alternative values. In Chomsky's image, it can be 'wired up' in two ways and, in the light of different sets of forms to which they are exposed, some children wire it one way and some the other. Suppose that the differences between two languages are wider. Then it may be that the principles include more variables. In the extreme case, the application of a principle in any form may be a parameter. In core language A it is effective in one form. In core language B it is wired up differently in one or two or more places. In core language C it is not effective at all; but that too can be one of the several different wirings that the universal grammar allows. We may also posit what in another context are called implicated universals (see Chapter 9, below). If a core language is wired in one way at point A, it can only be wired in such and such a way at point B. If principle A holds, principle B cannot hold; or if it does, a parameter X must have a particular value *q*, and so on. A child might then begin by fixing the value of some very basic parameter. Perhaps this is determined by some obvious property of the sentences to which it is exposed. The values of many others might then follow automatically, some so subtle that it might be hard to fix them directly.

A theory of universal grammar can thus incorporate a typology of

languages. In particular, Chomsky and his current wave of followers have talked of a 'configurational' and a 'non-configurational' type (traditionally, languages with fixed and free word order), or of 'pro-drop' languages (those in which a subject pronoun is used only for emphasis). But the motive is not to classify systems in a botanising fashion. Instead it is to explain how any internalised language, whatever its type, can develop. Once more, speakers can (according to Chomsky) make what, on the face of it, are inexplicable judgements. For example, he says that they can see 'with thought and preparation' that *John is too stubborn to expect anyone to talk to* is a sentence meaning that, because of his stubbornness, 'an arbitrary person would not expect' anyone to talk to John, whereas **John is too stubborn to visit anyone who talked to* is gibberish (Chomsky 1986: 11). They know such things 'without instruction or even direct evidence'; how then can they know them unless the relevant principles are innate? 'In many cases that have been carefully studied in recent work' (I am now citing Chomsky 1981: 3), 'it is a near certainty that fundamental properties of the attained grammars are radically underdetermined by evidence available to the language learner and must therefore be attributed to UG [universal grammar] itself'.

If one rejects this basic argument, the theory of core language will seem weak. On the one hand, a great deal of any internalised language lies outside it. The periphery will evidently include 'exceptions' such as irregular morphology or idioms (see again Chomsky 1986: 47). It will also include the idiosyncratic constructions taken by particular words. For example, *want* cannot take a *that*- clause (** The students want that Bill visit Paris*); however, 'we may assume' that this is 'an accidental gap reflecting properties that are not part of core grammar'. Although the construction is not 'idiomatic English', it is therefore 'fully grammatical at the relevant level of abstraction' (Chomsky 1981: 19). In this case, the counteraction of core grammar by the periphery may be trivial. But it is easy to imagine cases where it might be judged more serious. Suppose that a principle finds support over a wide range of languages; but, in just one, it does not hold. We could, of course, establish a parameter to cover this. But we might not wish to do so, since a universal grammar cannot be a mechanical accumulation of everything that we must posit in individual grammars. An alternative is to say that the principle does hold universally; but, outside the core, the speakers of this language must have internalised peripheral rules which (exceptionally) negate it.

On the other hand, the theory of core language can itself be weakened. Suppose that parameter A is, in general, set to q if parameter B is set to r; but, once more, there is one language where it is not. We could, of course, say that the implication is not part of universal grammar. But we might be reluctant, since any implication simplifies the child's task. An alternative is to say that it is given innately as the normal or, in Chomsky's terms, the UNMARKED case.

The language which is an exception represents a MARKED case where, in wiring up a core language, children will be compelled by the facts to set parameter B to r but A to some other value. We thus have a theory which incorporates alternatives, and can accommodate at least two sorts of exception. If it is considered simply as a theory of linguistic universals, there is a risk that it will be immune to counter-evidence.

But if one accepts Chomsky's basic argument, none of this is crucial. For, yet again, we posit that something is innate not because we have found it to be universal, but because we see no way by which it can be learned. In this discussion I have continued to use the word 'learn', as Chomsky too does occasionally. But a child's learning of a language is not seen as learning in an ordinary sense. Nor is it like a scholar's construction of a theory, as Chomsky had first suggested. In the mid 1970s, he compared the acquisition of 'cognitive structures' to the development of bodily organs (thus, in particular, Chomsky 1976: 10f.). Language itself was thus described, at first in inverted commas, as a 'mental organ' (36). It is not learned, but grows in the child's head by a complex interaction between genetically-determined structures and the environmental input through the senses. As Chomsky remarked later, the assumption that the mind has 'a rich innate structure' sits naturally with the belief that it is modular (1980: 40 ff.). Each module has its own innately structured properties, like an arm, an organ of vision, and so on.

At the time, a theory of the language organ barely existed. Its empirical study, now restricted to core language, is still highly idealised. Many, including the present writer, are not convinced that it exists. But as the result of many successive shifts in Chomsky's thinking, in which he has abandoned most of the ideas that had been central in his first phase, and a great deal of what seemed to be crucial in his second, he at least has a conceptual model that is appropriate to his ends.

REFERENCES

Bloomfield, L. (1933) *Language*, Holt, New York.
Bloomfield, L. (1934) Review of Havers, *Handbuch der erklärenden Syntax*, *Language*, 10: 32–9. Reprinted in Hockett 1970: 281–8.
Chomsky, N. (1957) *Syntactic Structures*, Mouton, The Hague.
Chomsky, N. (1959) Review of Skinner, *Verbal Behavior*, *Language*, 35: 26–58.
Chomsky, N. (1965a) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Mass.
Chomsky, N. (1965b) 'The formal nature of language', in Lenneberg, E., *Biological Foundations of Language*. Reprinted in Chomsky 1972a: 115–60.
Chomsky, N. (1966a) *Topics in the Theory of Generative Grammar*, Mouton, The Hague.
Chomsky, N. (1966b) *Cartesian Linguistics*, Harper & Row, New York.
Chomsky, N. (1968) *Language and Mind*, Harcourt Brace, New York.
Chomsky, N. (1972a) *Studies on Semantics in Generative Grammar*, Mouton, The Hague.