

## 6 Cognition, Environment and Language Learning

---

This chapter takes up some general issues concerning how linguistic knowledge develops. The first section reviews facts that argue that linguistic development is guided by a biologically given set of abilities. The following sections deal with questions concerning the acquisition in real time of linguistic knowledge, particularly the relative contributions of principles of universal grammar and of principles and mechanisms that lie outside grammar *per se*. The final sections of the chapter deal with the role of special speaking styles used to children and the relationship between language development and other aspects of cognitive development

### 6.1 Innateness

It is a standard assumption among linguists and many psychologists that much fundamental linguistic knowledge is innately given. That is, the child is assumed to be biologically equipped with knowledge of universal grammar – the basics of language structure. The child has, as it were, blueprints for all the possible types of language in her head (although not necessarily from the outset, as we will see below). In the course of language development she settles on the particular grammar of the language surrounding her.

Many facts about language development support in a general way the idea that there is a strong innate component of language structure. Language development exhibits many of the properties of biologically given behaviours (see Lenneberg 1967): three such properties in particular are worth noting here.

1. There is an *orderly progression of stages*. In different areas of grammar, we have seen that there is a distinct sequence in which children develop their linguistic abilities. For example, babbling emerges before first words; a one-word stage generally precedes the emergence of multi-word utterances; and specific morphemes and rules emerge in approximately the

same order across different children learning the same language (as, for example, in the order of emergence of English morphemes in Roger Brown's 1973 study). Although the age at which a particular stage is reached may vary considerably, there are rough guidelines (for example, the babbling period is generally placed between six months and the turn of the first year). A pattern of ordered stages in development is one of the hallmarks of biologically programmed behaviours (Lenneberg 1967). The existence of regular stages in language development is analogous to other biologically triggered phenomena in both humans and other species (such as walking in humans and flight in birds).

2. There is a *critical period*. Related to the fact that there are distinct stages in development is the fact that there is a critical age beyond which our ability to learn a language is significantly impaired. There is substantial debate about what exactly the critical age is, and what physiological changes it may reflect (see Lenneberg 1967; Krashen 1973; Curtiss 1977). None the less, from the early teens onwards most people exhibit a decreasing ability to learn a second language. And in those fortunately rare cases where an individual has been deprived of exposure to a first language in early and middle childhood, certain language skills may be unattainable (Curtiss 1977). Such observations support the view that early and middle childhood is a period in which we are biologically equipped to learn language in a way that we are not able to in later life. Again, there are analogies with other biologically programmed abilities. For example, kittens who are exposed to only restricted visual stimuli at an early period have long-term deficits in vision (see Kolb and Wishaw 1985, pp. 605–6 for a summary of pertinent research on vision).

3. Development is to a degree *independent of external stimuli*. Clear evidence of this is the fact that deaf children babble (see Locke 1983, ch. 1); such activity must be the result of a biologically timed programme that is not dependent on exposure to speech. Another case where a speech stimulus is either unnecessary or only minimally needed is categorical perception of speech sounds by infants who have had little exposure to speech (in chapter 2 we saw that six-week-old infants discriminate between speech sounds that adults categorize as distinct in a similarly categorical manner, and on a similarly fine-grained phonetic basis).

The existence of a biological programme for language learning does not preclude an important role for external stimuli. A child has to be exposed to language for normal development to take place. It is generally agreed that the child is an active learner, who 'works on' the speech she hears, using some form of grammar-forming mechanism. Current views of the nature of that mechanism have been formed in the context of several facts and assumptions about the nature of the learning situation. These are the topic of the next section.

## 6.2 Input and Errors

*Input* is usually understood to mean the speech forms to which the child is exposed, which may be augmented by contextual clues as to what an utterance means. The nature of the input itself provides a strong argument for the position that the child is equipped with a highly structured grammar-forming mechanism. The speech to which the child is exposed gives her only limited information concerning the correct rules for the language to be learned. Input can be divided into *positive* and *negative* evidence. The *positive* evidence (or positive input) is evidence that a particular form exists in a language. Hearing a sentence is positive evidence that it exists. One important limitation on positive input is that the sentences children hear do not contain overt information about their structure and meaning. In a sentence such as 'John told Bill to shave himself', the first two words must be analysed as subject noun phrase and verb, the reflexive pronoun 'himself' must refer to 'Bill' (it may not refer to 'John'), and so on. Sentences do not come ready tagged with information about what the component parts of the sentence are or what the sentence can or cannot mean, although the context of utterance may provide clues to possible meanings. Despite this lack of overt information about structure and meaning, children do manage to learn the rules of their language. Another limitation on positive input is that the child may hear only a limited sample of the sentence types that are actually grammatical in her language.

Not only do children get no overt information about structures and meanings, they are also not informed about which strings are ungrammatical. There is a lack of *negative* evidence. A child will generally hear only well-formed sentences and phrases (although the quality of the speech children hear has been a matter of dispute; see the final section of this chapter). Moreover, when children make errors in their own speech, they are not corrected by adults. In a well-known study, Brown and Hanlon (1970) found that parents responded to the truth value of children's utterances, but did not overtly correct ungrammatical forms produced by the child (see Demetras et al. 1986 for a more recent study making the same point). Despite the lack of negative evidence, the child manages to avoid or eliminate errors and arrive at the correct grammar of the language around her.

The restricted nature of the input (the paucity of positive evidence and the lack of negative evidence) is often referred to as the *poverty of the stimulus* and has frequently been invoked by Chomsky and others as a point in favour of a learning mechanism in which the child's innate knowledge of principles of grammar plays a major role in guiding development (see Chomsky 1959; Chomsky 1965; Hornstein and Lightfoot 1981). Without knowledge of such

principles, it is claimed, the input would simply be insufficient for the child to arrive at a system of the complexity of adult grammars. An example will help bring home this point.

Consider the set of sentences in (1):

- (1a) Who did John say that Sue kissed t?
- (1b) Who did John say  $\emptyset$  Sue kissed t?
- (1c) \*Who did John say that t kissed Sue?
- (1d) Who did John say  $\emptyset$  t kissed Sue?

In (1a,b) the object of the subordinate clause has been questioned; the question is grammatical regardless of whether the complementizer 'that' is present or not. In (1c,d) the subject of the subordinate clause has been questioned, and the question is only grammatical if the complementizer 'that' is absent. Although the matter is far from settled in linguistic theory, this superficially strange set of facts has been argued to follow from quite basic grammatical principles (and to be related to the phenomenon of null subjects, discussed briefly in chapter 4; see section 6.3 below). It is presumably knowledge of the grammatical principles involved that helps the child towards the knowledge that (1c) is an ungrammatical sentence in her language, since it is highly implausible that if the learner had to rely solely on mechanisms of hypothesis formation and generalization from the input she would arrive at the correct grammar. In fact, she might well be expected to arrive at an incorrect grammar (in which (1c) is grammatical).

In general the presence of 'that' before a complement clause is optional, a fact that the learner would be able to extract from examples such as

- (2a) John said that Sue kissed Bill
- (2b) John said  $\emptyset$  Sue kissed Bill

Moreover, the evidence of simple questions (in addition to (1a,b,d)) will tell her that both the subject and object positions can be questioned:

- (3a) Who will arrest him?
- (3b) Who will he arrest?

The learner working with minimal powers of hypothesis formation based on the sentence types she may actually hear might be expected to form a grammar that includes the specification 'the complementizer "that" is optional' and a

question formation rule that questions subjects as well as objects. To put it another way, the learner would be predicted, without access to the constraint that blocks (1c), to form a grammar in which (1c) is grammatical, since that is the simplest grammar consistent with the speech forms she hears. This example is a good illustration of the 'poverty of the stimulus'; the learner will have no overt information as to the ungrammaticality of (1c) and the sentence types she may hear (the positive evidence, such as the remaining sentence types in (1) and the sentence types in (2) and (3)), are fully compatible with an incorrect analysis, in which (1c) is a grammatical sentence. The fact that children are capable of ending up with grammars that preclude questions of the type in (1c) is *prima facie* evidence that the child has access to whatever principles of grammar determine the adult grammaticality facts; it is highly unlikely that the child could deduce the facts from the data only.

Two other observations go hand in hand with the limited nature of the input in arguing for a highly structured learning mechanism. These are that language learning is a relatively error-free process and that language learning is a fairly rapid process. Both of these claims can be challenged to a degree. Errors may go unnoticed, and some learning (as we have seen in chapter 4) may go on well into the school years. But none the less the impression left by observation of child language is that the errors children make are of a limited kind and that by five years and plausibly much younger children have a linguistic system that conforms in essential ways to that of the language they are learning.

### 6.3 The Role of Universal Grammar in Language Development

We will assume that the child analyses input sentences into strings of words via a sentence processing mechanism. (The exact nature of the processing mechanism – for both children and adults – is a complex and vexed issue, and is the topic of the next chapter). Such strings will then be analysed by existing rules of the language being learned and/or principles of universal grammar. The primary role of universal grammar in language development is to limit the hypotheses that a child can form concerning the rules of his language, thus also limiting errors and helping explain the speed and ease with which language is learned.

An important aspect of the idea that universal grammar guides language development is the fact that use of principles of grammar potentially allows the child to form rules and hypotheses that affect and determine the form of

sentence types for which he may have no direct evidence in the input itself. For example, sentences such as (1c) above are not ungrammatical in all languages. In Italian, the equivalent of (1c) is grammatical. This fact has been argued to be linked in a principled way to the existence of null subject sentences (see Rizzi 1982). We will not detail the theory here; the analysis is not uncontroversial, nor is it easily extended to all languages that allow sentences such as (1c). The basic argument runs as follows: sentences such as (1c) in English are ungrammatical because a principle of grammar (the 'empty category principle') is violated by movement from the subject position if the complementizer is present. Italian allows the subject and verb to be inverted in non-interrogative sentences (so one can have sentences equivalent to 'Hit Mary John' meaning 'Mary hit John'); it also has null subject sentences (the equivalent of 'walks' for 'she/he/it walks') and a special type of (unpronounced) pronoun will occupy the subject position in null subject sentences and sentences with inversion of the order of subject and verb. Sentences of the type (1c) will be permitted because movement is possible from post-verbal position in the inverted sentence form. Regardless of the ultimate correctness of this analysis, the basic point concerning the role of universal grammar is plain. Knowledge of the grammar of very simple sentences (whether or not the language permits inversion of the subject and verb, and whether or not the language has null subjects) can allow the child, if he is equipped with the appropriate principles of grammar, to make inferences concerning quite complex sentence types about which he may have no direct evidence. One type of sentence can act as a 'trigger' for the formation of rules that govern other types of sentence. Thus, while we argued in the previous section that general inference mechanisms unconstrained by principles of universal grammar would lead the child on the basis of simple sentence types to conclude that (1c) is grammatical in English, if the mechanism for forming grammatical rules is governed by principles of grammar simple sentence types of the language at hand can promote the formation of grammars with rules that make correct predictions concerning the grammaticality of sentences the child has not heard.

## 6.4 Learnability and Acquisition Principles

*Learnability theory* is concerned explicitly with the conditions under which successful learning of a system of rules can take place within a finite amount of time. The terms *learnability* and *learnability theory* were originally associated with work on learning formal languages, which may conform only to a

degree with natural language systems (Gold 1967; Wexler and Culicover 1980; Wexler 1981). But they have come to be used in a more general sense, to refer to the study of conditions that will permit successful language learning in a limited time span. In this more general sense, learnability theory takes in questions such as the relative contribution of principles of universal grammar and input (the speech forms the child hears) in rule formation, the prevention of errors in learning and the correction of errors that have been made. We consider here two learnability proposals, the first designed to help explain the speed with which language is acquired, the second designed to prevent the learner from making errors.

#### 6.4.1 *Subjacency and Degree-n Learnability*

In chapter 4 we listed some structural positions from which a question word may not be moved, including relative clauses, temporal clauses and indirect questions:

- (4a) \*What did John see a horse that kicked?  
(cf. John saw a horse that kicked a box)
- (4b) \*What did John read Dickens before writing?  
(cf. John read Dickens before writing that memo)
- (4c) \*What did John wonder who loved?  
(cf. John wondered who loved egg rolls)

We saw that there is evidence that from an early stage children learning English are obedient to the constraint on extraction from a temporal clause, and show some sensitivity to the other constraints at least by age five.

Does the existence of such constraints aid the language learner? It can be argued that such constraints do aid learning, in the following sense. Although (as we saw in chapter 4) there are apparent cases of violation of at least some of the constraints, cases where the constraint is not followed are easy to spot, in that they occur in sentences of limited complexity. For example, one finds languages where the equivalent of (4a) is grammatical, but one does not find languages where (4a) is ungrammatical but (4a') is grammatical:

- (4a') What did John see a horse that kicked a box that contained?

That is, if a language is going to allow sentences that allow questioning into a relative clause, it will allow such questioning in the simplest possible structures containing a relative – it will not be the case that such questions are permitted only if a relative is embedded two sentences down, or four sentences

down, etc. The same is true of the other constraints mentioned above. (Recall from the discussion in chapter 4 that apparent violations of the constraints on movement are frequently assumed to follow from formation of the offending sentences by a means other than movement.)

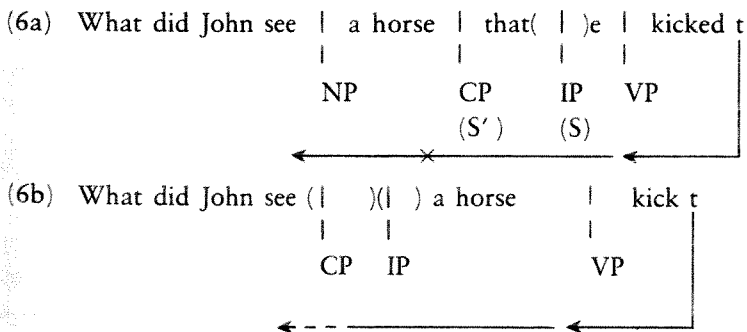
The facts will be accounted for if the grammatical principles that determine the ungrammaticality of the sentences in (4) are such that they must show up on the simplest sentences to which a constraint could potentially apply – i.e. on the first level of embedding. The *subjacency* principle is one component of the principles proposed to account for the constraints on movement rules and it has precisely the effect of being formulated so that it will restrict movement from relatively simple structures. The principle, which originates with Chomsky (1973), can be stated roughly as follows:

*The subjacency principle:* an element can move only over one layer of structure (i.e. it can move only from a layer that is *subordinate* and *adjacent*)

The correct definition of ‘layer of structure’ is a vexed question. It will be sufficient here to illustrate the role of subjacency in accounting for the difference between the ungrammatical (4a) and a grammatical sentence such as (5), where movement has taken place from within the complement to the verb ‘see’, rather than a relative clause:

(5) What did John see a horse kick?

In the framework of Chomsky (1986b), all phrasal nodes can potentially constitute (by various mechanisms) layers of structure pertinent to the computation of subjacency values in movement. A phrase has the potential to move from a position infinitely far down in a structure by crossing no more than one layer at a time (in accord with subjacency). The difference between (4a) and (5) is that (4a) contains a stretch of structure ‘the horse that’ with two layers that must be crossed, in violation of subjacency:



(The IP layer in (6a) and the CP and IP layers in (6b) are parenthesized because other factors void them; in most analyses, the subject position of the relative in (6a) is also derived by movement, not shown in (6a).)

Without spelling out more details, this account may seem arbitrary; however, the abiding virtue of the notion of subjacency in grammatical theory is that it offers a unified or semi-unified account of various constraints. On any version of the formulation of subjacency, it refers to phrasal nodes. In the version of Chomsky (1986b), all phrasal nodes are pertinent, subject to various principles and conditions; in other versions, particular nodes (NP, S/S') are singled out as pertinent to computing subjacency values. All of these node types are present and embedded one within the other in D-structures with minimal levels of embedding; hence the complexity of sentences needed to see if subjacency is not operative (i.e. to see that the construction is *not* subject to the constraints that characterize movement) is correspondingly limited. The subjacency principle, in combination with the nature of the phrase structure system of the language, ensures this simplicity of necessary input data.

To use terminology deriving from Wexler and Culicover (1980), grammatical constructions are degree- $n$  learnable, where  $n$  corresponds to a numerical measure of complexity of structures needed to deduce the adult grammar from the input data. Following the custom in discussions of learnability, we can take the number of embedded sentences to be the measure and say that, given that the learner works with knowledge of subjacency and that subjacency constrains movement rules, whether or not a particular construction is derived by movement will be degree-1 learnable, that is, whether movement is involved will be deducible on the basis of structures with no more than one embedded sentence.<sup>2</sup>

#### 6.4.2 *The Subset Principle*

In the case of examples such as the one in the previous section, learnability theory does little more than articulate the advantage to the learner of knowledge of universal grammar. If the learner has prior knowledge of the subjacency principle and the fact that the principle constrains a particular type of grammatical operation (movement between D-structure and S-structure), the range of hypotheses he has to entertain with respect to his language is correspondingly limited and language learning should be made easier. The fact that subjacency limits the complexity of sentences that the learner will need to hear contributes to an explanation that language is learned in a finite amount of time.

But learnability theory has also thrown up hypotheses that are not in any obvious way a reflection of principles that govern adult grammars. The *subset principle* (Berwick 1985; Wexler and Manzini 1987) is an example. This is

a principle designed to prevent the learner falling into error when more than one possible analysis is permitted under principles of grammar. As such, it is a principle of language acquisition, rather than a principle of grammar. Its potential application can be illustrated with respect to the binding of pronominal elements.

We saw in chapter 4 that principles A and B of the binding theory ensure that reflexive pronouns and definite pronouns are in approximately complementary distribution in English. Principle A requires that a reflexive be bound in its local domain; principle B requires that a pronoun be free in its local domain. Taking the local domain to be the first sentence node above the pronominal element, we saw that the principles ensure that 'himself' refers to 'Fred' in (7a), and not to 'Tom'; and, conversely, that 'him' may refer to 'Tom' but not to 'Fred' in (7b):

(7a) Tom said [that Fred had shaved himself]

(7b) Tom said [that Fred has shaved him]

The definition of domain is not invariant across languages, nor is it such that there is complete complementary distribution between reflexives and definite pronouns within a language (see, for example, Wexler and Manzini 1987). Considering just the case of reflexives, there are languages in which a reflexive may refer to the subject of a higher sentence. In Icelandic, for example, a reflexive may refer to a higher subject provided the lower clause is not tensed, and in Korean a reflexive may refer to the subject of a higher clause even if the lower clause is tensed. In Korean, therefore, the equivalent of a sentence such as (7a) will be ambiguous; the reflexive could refer to either the higher or the lower subject. Plainly, since both possibilities (reference to only the subordinate clause subject or reference to either the main or subordinate subject) are allowed in human languages, principles of universal grammar will not give a unique answer as to what the rule is, for any string of words superficially equivalent to (7a). The subset principle (see Berwick and Weinberg 1984; Berwick 1985; Wexler and Manzini 1987) says roughly that in such a situation the language learner opts for the grammatical system that is least permissive – the system that yields the smallest range of grammatical sentences.

The prediction of the subset principle with respect to the example above is that the child learning Korean will start off with an English-type grammar, in which the reflexive can be bound only to the lower subject. Only at a later point will the child revise her grammar to the more permissive, Korean-type system. This should not be a difficult matter, assuming the learner will eventually hear sentences where there is evidence from the context and/or the internal structure of the sentence that the more restrictive rule system is incorrect (for example, the learner might encounter the equivalent of the ungrammatical English sentence 'Tom said that Linda shaved himself').

The appeal of the subset principle in learnability terms is that it promotes error-free learning and so minimizes the need for negative evidence (evidence that a particular form is ungrammatical in the language to be learned). If the child learning English erroneously decided that the reflexive pronoun could refer to the higher subject in sentences such as (7a), how would the error ever be discovered?

There are at least two criteria by which we can evaluate any proposed principle of acquisition, such as the subset principle: (1) Does the principle translate into plausible mental mechanisms? (2) Do the facts of language learning correspond to the predictions of the principle? To the extent that there is evidence available, the performance of the subset principle with respect to these criteria is mixed. With respect to the first point, it is not clear what it might mean to say that the learner selects the most restrictive grammar. Is one to suppose that the learner actively compares the possible analyses, somehow unconsciously reasoning along the lines: 'Heavens, I've no evidence that binding a reflexive to a higher subject is allowed, I'd better stick to lower domain binding for the time being' (see Fodor and Crain 1987 for comments to this effect)? If so, why does similar reasoning not prevent the other overgeneralizations and innovations that the learner does make (for example, the use of innovative causative forms described in chapter 3)?

There is at least some empirical evidence that the predictions of the subset principle are correct, in terms of children's performance. Lee (1987) reports that children learning Korean do indeed prefer the structurally closest possible referent for a reflexive pronoun, in contrast to adult speakers of Korean, who more frequently select a structurally distant referent for a reflexive. (However, not all the recent studies of knowledge of reflexive binding yield a pattern consistent with use of the subset principle; for example, Jakubowicz and Olsen (1988) report data difficult to account for under the subset principle).

The subset principle is an example of a principle that appears in essence to be a principle of learnability – one which provides the learner with an orderly procedure for positing hypotheses about the structure of the language. The principle will prevent errors and so help explain the fact that language is learned in a limited time span.

## 6.5 Summary: Components of a Learning Model

We can summarize the points made in the previous sections with a diagram (see figure 6.1). The input to the learning mechanism is analysed by a sentence

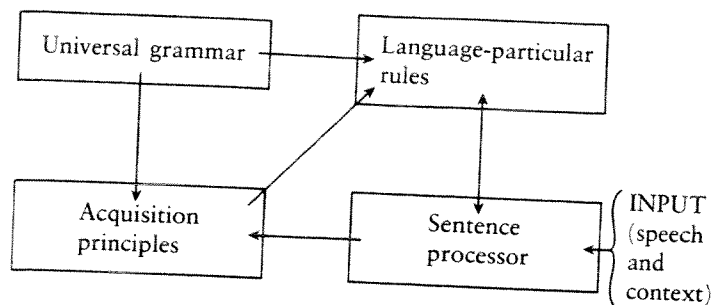


Figure 6.1 Components of a learning model

processor; if the learner's current language-particular grammar contains no rules suitable for the input, the input string may be subjected to analysis by acquisition principles, whose operation is, like language-particular rules, constrained by universal grammar.

## 6.6 Some Questions and Problems in Acquisition Theory

The roles of universal grammar and of acquisition principles are quite easy to describe in general terms: universal grammar provides the child with analyses of the data; acquisition principles help resolve remaining ambiguities. But there are also many substantial questions and grey areas with respect to the contribution of universal grammar and candidate acquisition principles in real-time learning. Here we briefly consider three related questions: Does learning follow sequences dictated by markedness values? Do children ever violate principles of grammar? And is 'parameter setting' as an account of learning a significant advance over hypothesis testing, within a framework where universal grammar governs children's hypotheses?

### 6.6.1 Markedness and Orders of Acquisition

In chapter 2 on phonology, we saw that the order in which sounds emerge in child speech is to a degree predictable in terms of the frequency of the sounds cross-linguistically. Sounds that are very common in the world's languages tend to emerge earlier than rarer sounds. In the terminology introduced there, unmarked sounds emerge before marked sounds. We can ask in more general

terms whether normal, frequently occurring constructions and rules correspond to the earliest forms and rules in children's grammars. As noted in chapter 2, one goal of linguistic theory is to account for distributional facts (whether a form is frequent or rare cross-linguistically) in terms of principles of grammar. Early emergence of unmarked forms is thus potentially evidence for the guiding force of principles of universal grammar.

As far as the evidence goes, it is probably correct to say that unmarked rules and forms are early acquired. The facts of phonological acquisition are largely consistent with that claim and so are some studies of the development of syntactic constructions. To take one example, in English verbs such as 'give' permit two, more or less equivalent, syntactic means of realizing the semantic roles *theme* (object transferred) and *goal* (recipient) in the verb phrase, as illustrated in (8a,b):

- (8a) John gave the book to the school  
 (8b) John gave the school the book

In (8a), the theme ('the book') is direct object and the goal ('the school') is the object of the preposition 'to'. In (8b), there are two NPs in the verb phrase, with the first having the goal role and the second the theme role. It is generally agreed that the NP PP form in (8a) is the unmarked or typical case: that syntactic frame can occur freely with verbs with a suitable basic meaning, whereas the distribution of the NP NP form (8b) is more restricted. (For example, if we substitute the verb 'donate' for 'give' the result is grammatical for (8a), but not for (8b).) Thus we can consider the NP PP form as normal or unmarked, and the NP NP form as marked, based on distributional facts; moreover, a principled theoretical account of these 'markedness values' is available (Stowell 1981). For these constructions, the facts of language acquisition fit fairly well with the markedness values based on distribution (and linguistic theory). Cook (1976) and Roeper et al. (1981) both find NP NP forms to be harder for children to comprehend than NP PP forms, consistent with the view that the NP PP form is the unmarked construction.

Although the facts of acquisition sequences for syntax as well as phonology can be fitted to observations about marked and unmarked forms, the evidence is perhaps less clear-cut in the case of syntactic development, where the role of markedness values may be obscured by other factors. For example, although the experiments cited above have shown that children do better at comprehending NP PP dative sentences than they do at comprehending NP NP datives, NP NP forms with the verb 'give' do occur in young children's speech quite early on. Brown (1973), for example, provides the following example from a two-year-old:

- (9) Give doggie paper

Plausibly the high frequency of NP NP forms with the verb 'give' in the speech the child hears promotes its early use by the child. (The role of input in learning such structures is returned to below.) Thus, although the experimental studies may reflect some kind of basic asymmetry between the two forms, the child may early on have some mastery of a marked form for reasons external to the structure of the grammatical system.

One interesting question concerning sequences of development is the relationship between markedness and candidate acquisition principles. Cross-linguistic and language-internal frequency is the standard rule of thumb for unmarked situations. The subset principle as currently formulated is not a principle of grammar that is 'active' in grammars of adult languages. But the subset principle does predict that the first-acquired rule should be the most frequent rule cross-linguistically. The logic of the situation is as follows. If there exist languages with only A (where A is some sentence type permitted by rules of grammar – in the example discussed in section 6.4.2, binding of a reflexive to a lower subject) and languages with both A and B (where B is some additional sentence type – in the example, binding to a higher subject), but not languages with B only, it must be the case that A is the most frequent situation in languages of the world, regardless of which language type (languages with A only or languages with A and B) is more common in terms of the distribution of actual human languages. In other words, the logic of the subset principle dictates that the first rules the language learner hypothesizes are the most frequent rules (situations) cross-linguistically. So the predictions of the subset principle are in a manner confounded with markedness predictions. Thus it is possible that the subset principle, at present formulated as a principle of acquisition rather than a principle of grammar, may ultimately reduce to (as yet undiscovered) rules of grammar; or, conversely, that frequent situations in languages of the world do not always follow from principles of grammar, but may reflect in some manner the process of acquisition.

### 6.6.2 *Continuity or Maturation?*

The first section of this chapter summarized some general arguments in favour of the view that language development is a biologically guided behaviour. However, a strong component of innate linguistic knowledge does not entail that all constraints and principles of universal grammar are available to the child from the outset. It is logically possible that some grammatical rules and constraints are programmed to emerge (under appropriate conditions of

external stimulation) only after a period in which they are not present in the child system. It is not difficult to think of analogies with other aspects of human development – for example, the physiological changes of puberty are biologically programmed changes whose exact timing will depend on conditions in the individual's environment, such as diet.

Although there are many possible ways to frame the problem in detail, we can identify three broad positions about whether some principles of grammar emerge only after a period in which they are absent from children's grammars (this taxonomy is taken from Weissenborn et al. 1992):

1. *Strong continuity.* All principles and constructs of universal grammar are available at the outset and each grammar formed by the child is a correct (partial) grammar for the language to which the child is exposed.

2. *Weak continuity.* All principles and constructs of universal grammar are available at the outset and all child grammars will be 'possible human grammars', in the sense of falling within the patterns of adult grammars (either observed or permitted under the theory). The child's grammar may, however, deviate from that of the language he will ultimately acquire.

3. *Maturation.* Some properties of universal grammar mature. That is, some properties of grammar are biologically programmed to emerge only after a certain period of development. If such a property is an absolute universal (i.e. holds obligatorily for structures and rules to which it is relevant) then child grammars may of necessity fall outside the range of 'possible human languages'.

Since the 1970s, the most popular position in the literature has been some version of weak continuity. That position is appealing, allowing for the child to form rules that are not present in the language he hears, but none the less constrain the acquisition process. However, recent studies have proposed maturational analyses for some grammatical phenomena; see Felix (1987) and Borer and Wexler (1987).

At the present state of evidence, the data in support of maturation do not convincingly challenge the weak continuity position, since the facts attributed to the maturation of grammatical principles are generally amenable to alternative analysis, compatible with strong or weak continuity. For example, Borer and Wexler suggest that the acquisition of the passive construction may follow from maturation of the ability to move an NP into the subject position. As outlined in chapter 4, passives are formed in English by movement of a direct object into an empty subject position:

- (10) [e] were eaten the bananas  
The bananas<sub>i</sub> were eaten [t<sub>i</sub>]

The suggestion that children do not form passives by movement does not mandate a maturational account of the child's development, provided that the theory of grammar gives some alternative way for children who have not undergone the proposed maturational change to form passives without movement. We saw in chapter 4 that there is such an alternative, with the 'passivized' subject NP in subject position from the outset.

Lack of clear counter-evidence does not entail that the continuity hypothesis is right and maturational accounts are wrong. Very early speech (one-word and telegraphic utterances) is a tempting candidate for grammatical systems which are genuinely impoverished by comparison to adult grammar and the first two years of grammatical development at least may involve maturational changes that bring the child's grammar into line with the general shape of adult grammars (see chapter 4, section 4.6). Stevenson (1992) and Clahsen (1992) both discuss the complexities of identifying maturational developments in grammar.

### ***6.6.3 Parameter Setting vs. Hypothesis Testing***

Another topic that has been at the forefront of recent discussions on the role of universal grammar in language development is whether there is an advantage to 'parameter setting' as a model of development.

An elementary example of parameter setting examined in chapter 4 was the order of heads and modifying material in phrase structure. A language such as English fairly consistently puts heads of phrases to the left; a language such as Japanese fairly consistently puts heads of phrases to the right. So in English a noun precedes a complement sentence or relative clause, a verb precedes its object(s) and complement sentences, etc., whereas in Japanese the general pattern is reversed, with a relative clause preceding its head noun, etc. This division in language types can be accounted for in terms of a general 'head position' parameter for phrase structure, which is set to one value for English and to another value for Japanese. The need for many individual rules (one for each phrase type in the language) is thus in principle obviated and learning made easier: once the child has worked out what the pattern is for one type of phrase, she can then automatically project the correct pattern for other phrase types in the language.

Learning as 'parameter setting' has been opposed to learning as 'hypothesis testing'. On the latter view, the child would have to sort through and decide about a range of hypotheses that she never need entertain under the parameter setting model. To continue the example of basic phrase structure, the child who is equipped with knowledge about what a phrase structure can look like, but not with the knowledge that there is a contingency relationship between

the existence of one phrase structure and others, will have to decide for each phrase type separately (NP, VP, PP, etc.) what the shape of the phrase is. So the learner who has knowledge that the head-position parameter imposes regularity across the shape of phrases in the language will only have to make one decision (is her language head-first or head-final?), but the learner who does not have knowledge of the parameter must make several separate decisions (is the noun phrase in her language head-first? is the verb phrase in her language head-first? etc.). Since one decision should be easier than many, it is argued that the parameter setting model helps explain the ease with which children learn language.

Despite the intuitive appeal of this argument, the differences between learning under a parameterized view of grammatical systems and learning as it was conceived under previous models of generative grammar may not be very large. First, although the terms parameter setting and hypothesis testing are often used as if they represented two different styles or methods of learning, they do not in fact represent a difference in the learning process itself. Parameter setting *is* hypothesis testing: the child must, for example, test the hypothesis that her language is either head-first or head-final against the speech forms to which she is exposed. All that differs potentially is the *number* of hypotheses that have to be checked out; parameters that define clear-cut contingency relationships between grammatical phenomena reduce the range of hypotheses to be tested. The potential learning advantage in parameter setting depends, then, on how clear-cut the contingency relationships specified by the parameters are. Very often the contingencies are complex and imperfect, once more than the most elementary examples are considered (see Atkinson 1987; Bach 1988; Roeper and Weissenborn 1990; Valian 1989 for pertinent discussion).

## 6.7 The Limits of the Linguistic Model: Lexical Learning

In this section we look at the role of mechanisms of learning that lie outside either universal grammar or candidate principles of acquisition. In a restricted range of cases, it seems plausible to assume that acquisition of the details of grammar depends on distributional features of the input, such as the frequency with which particular lexical items are encountered in particular syntactic environments.

The idea that input and frequency constitute an explanation of anything is something of an anathema to many generative grammarians (though not N.