Language acquisition and the architecture of the language faculty

Peter W. Culicover

The Ohio State University

Language acquisition and the architecture of the language faculty

Peter W. Culicover
The Ohio State University

1.   Introduction

This paper explores the relationships between syntactic theory, the theory of language acquisition, and the course of language acquisition, viewed from a cognitive science perspective. The main questions that I will address are these:

C   In what ways do considerations about language acquisition constrain the architecture of syntactic theory?
C   What is the relationship between a syntactic theory and what is actually in the mind/brain of a speaker?
C   In what ways can a syntactic theory explain aspects of the course of syntactic development?

I argue that the kinds of grammars that we are familiar with are not models of what is in the head; they are descriptions of its **behavior** to a certain degree of precision. So the course of development of this device cannot really be explained by the formal properties of the descriptive apparatus. We need to give some explicit consideration to the architecture of the language faculty itself, and we need to be clear about the status of the grammatical description in our theory of the architecture. Of course, to the extent that a description of the sound/meaning correspondences of a language is a correct one, the description contributes to a fuller understanding of the nature of the learner's task.

2.   Grammars

2.1.   Views of grammar

Let us first briefly reconsider some traditional views of grammar. There are essentially two traditional views. On what I will call the Cognitivist view, a grammar is an idealized description of something that is in the head of a native speaker of a language, which we may call the Mental Grammar. It is assumed that there is some capacity for language, called the Language Faculty, that produces a Mental Grammar given linguistic experience (Chomsky 1965). On this view there is a certain correspondence between the form of the grammatical description and the Mental Grammar. The architecture of the grammar is to some extent a claim about the architecture of the Language Faculty. An adequate grammar on this view is one that correctly captures some property of the Language Faculty.

For example, if the grammar categorizes words into Nouns, Verbs, etc., and posits functional categories such as Tense, Agr, and Aspect, and incorporates a constraint such as Subjacency or the Minimal Link Condition, these are all taken to be more than descriptive conveniences; they are assumed to be part of the Language Faculty itself. They make their appearance in a particular grammatical description as a consequence of this fact.

The other view is that the grammar is simply a description of the observed stuff of language, and bears no relationship to the Mental Grammar. On this Descriptivist view a grammar is at best an accurate description of what is possible in a language, and bears only accidentally if at all on the architecture of the

Language Faculty. An adequate grammar on this view is one that is maximally compact and perhaps satisfies certain formal or computational criteria.

These different views reflect Chomsky's distinction between I-language and E-language (Chomsky 1975). I should stress that there is no necessary correlation between a particular approach to grammar and whether it is Cognitivist or Descriptivist. That is, there is nothing inherently Cognitivist or Descriptivist about theories such as Principles and Parameters and HPSG. It is a question of what status we wish to claim for our descriptions.

There is an intermediate view, which I will call the Dynamical view. On the Dynamical view, the Mental Grammar is a dynamical system that passes through a set of states in a specific temporal sequence (Culicover 1999a). As it passes through these states it produces and/or processes sounds and meanings. The set of possible sequences of states is the set of linguistic expressions in the language embodied in this dynamical system, and the grammar is the description to some degree of precision of the behavior of this system. Hence, the grammar is not a description of the system itself, but corresponds to the architecture of the system to the extent that it correctly characterizes the properties of the behaviors that the system produces, abstracting away from practical limitations such as memory, etc. (Chomsky 1965). Working out the details of such a view is far from trivial, if it is even possible, and I do not wish to suggest that there are presently significant results that argue strongly in favor of it. Nevertheless, I believe that the conceptual motivations are appealing enough to take it seriously, and the empirical motivations discussed below are suggestive.

2.2.   "Capturing" things

On the Cognitivist view, capturing generalizations has traditionally been of primary importance. This stems from the belief that generalizations are made by a language learner in the course of language acquisition. Thus they are psychologically real (i.e. incorporated into the Mental Grammar) and must be represented in some way in the grammar. The priority given to such generalizations is reflected in the core-periphery distinction, where the core is understood to correspond to Universal Grammar (i.e. the Language Faculty), while the periphery is taken to be the product of poorly understood mechanisms of analogy, mutation, noise, and other deviations from perfection. The point is often made that it is essential to distinguish between generalization and idiosyncrasies in the grammar so that the grammar can formally reflect the fact that certain phenomena are characteristic of human language (hence reflect the architecture of the Language Faculty) while others are not. I will have more to say about this shortly.

On the Descriptivist view, what must be captured are the just facts. Since there is no sense in which the relative generality of the description corresponds to or is a claim about the architecture of the Language Faculty, parsimony and generality are primarily aesthetic considerations, a corollary of general scientific injunctions that prefer simplicity of description, other things being equal

The Dynamical view combines the two. On this view, all linguistic experience that leads to learning is realized in the Mental Grammar in some way, whether it is a generalization like "V precedes NP" or something very idiosyncratic, like the meaning of a particular lexical item or the properties of an idiom. Phenomena that are more general are presumably representative of more general properties of the human mind, but even the most idiosyncratic of linguistic facts represents a possibility of the human mind. I suggest that both wide-spread generalizations and isolated idiosyncrasies are symptomatic of linguistic universals. The former clearly are, but the latter are also, in that they reflect the universal capacity of the human mind to deal with linguistic phenomena of perhaps arbitrary complexity and specificity.

Let us turn next to the question of what sorts of information a grammar must incorporate.

3.   The Mental Grammar

3.1.   Parameters

Now, in considering the question of what is in a Mental Grammar and how it gets there, we should take note of the notion of parameter, since it is closely tied to the core-periphery distinction and the question of what a grammar is a description of. The idea of a parameter has two parts. One is that there is a finite number of dimensions on which languages may vary. The other is that on each dimension there is a finite, preferably small number of values that a language may choose among. A **deep** theory of parameters is one in which the parameters bear on relatively abstract properties of the grammar and the parametric values are a **principled** subset of the set of logically possible alternatives. Such a theory must be distinguished from a **superficial** theory of parameters in which the word "parameter" is used simply to refer to observed differences between languages.

The relationship between language acquisition and linguistic theory, in this framework, is that the values of the parameters are set by a learner on the basis of linguistic experience. If the theory of parameters turns out to be a theory of superficial parameters, then this perspective will not prove to be particularly explanatory, since it basically says that all of the properties of a language are identified by a learner on the basis of linguistic experience and incorporated into the Mental Grammar. But the theory of language acquisition is more interesting on a theory of deep parameters. Here the claim is that parameter setting is triggered by experience but there are certain facts about a language that follow immediately from the parameter settings that may not be identified by a learner in the data. They follow because of the fact that certain possibilities are excluded once the parameter values are determined.

For example, suppose that there is a principled connection between V-final and non-*wh*-movement, explained in terms of some abstract grammatical mechanism. It would be sufficient for a learner to determine that a language is V-final; it would then know that it could not have *wh*-movement, even prior to any linguistic experience that would be sufficient to establish whether *wh*-movement is possible in the language. One might expect a learner to be able to process and produce *wh*-questions with *wh*-movement prior to such experience.

3.2.   Correspondences

Viewing matters in the most general possible terms, the sound/meaning relation is one of **correspondence** as characterized by Jackendoff 1983. Either the grammar expresses the correspondences explicitly, or the correspondences fall out of more complex interactions between various levels of representation. I adopt here the general approach taken by Jackendoff 1997. One of the important properties of this approach is that the vocabulary in which to express the two sides of correspondences is not constrained. It is possible, therefore, to state correspondences involving single lexical items, phrases, classes of lexical items, classes of phrases, and more complex expressions. On the other side, the conceptual structures may be of arbitrary complexity, but are formed from a set of universal elements according to fixed rules of combination. Although it might appear that such a system has too much expressive power, Jackendoff argues that the power is needed in order to express the full range of correspondences that actually occur in natural language. For example, a lexical entry is a correspondence between a single word and a meaning. The simplest case would be one like the following. For ease of presentation I put the orthography between "/" to stand for the phonological form.

(1)   *dog*: </dog/, N, DOG>

The word *dog* is a noun that expresses the concept DOG.   More complex cases involve multi-word expressions,

(2)   *kick the bucket:* </kick the bucket/,V,DIE>

more complex conceptual structures,

(3)   *paint*: </paint/,V, SPREAD PAINT ON>

or both. The framework is capable of accommodating idiosyncratic expressions such as proverbs, and of associating aspects of conceptual structure with particular properties of syntactic constructions, as in (4). (Here I use "X" as a variable that ranges over verbs.)

(4)   *X one's way*: </X one's way/, V, MAKE PROGRESS BY X-ING>

A similar approach is taken in Construction Grammar (e.g. Goldberg 1995).

The conventional view of syntax is that it is the interface between sound and meaning, that is, phonology and conceptual structure. We may think of parametric variation as consisting at least in part of variation among different ways of expressing aspects of conceptual structure, on the view that the latter is universal. For example, the parameter that incorporates the V-initial/V-final property of VPs characterizes at least a portion of the variation in the ways in which languages encode the predicate, which we would assume to be universal in conceptual structure.

One important consequence of thinking of the sound/meaning correspondence in terms of correspondence rules is that it is not naturally compatible with a theory of deep parameters, since it does not take a subset of syntactic phenomena as counting as 'core grammar,' falling under the syntactic theory, while the rest falls outside of the theory, in the 'periphery.' On the correspondence view, the correspondences fall along a spectrum, from the most general to the most idiosyncratic. There are, of course, very general correspondences, which are those that have been the province of the theory of core grammar, e.g., phrase structure, exceptionless rules such as wh-fronting and sc rambling, syntactic categories that are universal or virtually universal, such as noun and verb, and so on.

In contrast, a theory that distinguishes core grammar from everything else must allow for the irreducible idiosyncrasy of the lexicon, but otherwise seeks to focus on those aspects of language where it appears that it is possible to capture generalizations. As noted, the motivation for this focus on the most general phenomena is the intuition that the general reveals what is universal in language and thereby tells us something about a fundamental aspect of human cognition, the capacity for language, while the idiosyncratic does not.

On closer examination this intuition appears somewhat flawed. Since human beings acquire both the general and the idiosyncratic, there must be a mechanism or mechanisms that can accommodate both. The mechanism for dealing with the idiosyncratic in language is every bit as unique to the human species as is the mechanism for dealing with the general. If there is in fact a spectrum of generality in the sound/meaning correspondences, as Jackendoff suggests, then to stipulate that only the maximally general is of interest to linguistic theory and to the study of the mind is to decide a priori what should be a matter of empirical inquiry. Even if we assume that the most general correspondences are instantiations of linguistic universals that permit only simple parametric variation, the question of how the rest of linguistic knowledge is acquired is left completely unexplored.

The variety of correspondences can be formulated in terms of structural descriptions stated over the

syntactic structures of a language. The complexity of the correspondences can be expressed in terms of these structural descriptions, and in terms of the mappings between these descriptions and the corresponding conceptual structures. Let the general form of a correspondence be

(5)   SS
        ∧
        CS

Where both SS and CS are labeled bracketings. The most general correspondences are those for which the SS is a minimal structure, e.g.

(6)   [$_{XP}$ X YP]

and the corresponding conceptual structure is isomorphic to it, e.g.

(7)   [$_{XP}$ X YP]
        ∧
        [$_{XP'}$ X'(YP')]

    Correspondences that are more complex may still be very general, but require specific information about morphological form or linear order. For example, Culicover and Jackendoff 1999 argue that the correspondence for the conditional comparative construction illustrated in (8) is (9).

(8)   The more it rains the angrier I get.

(9)
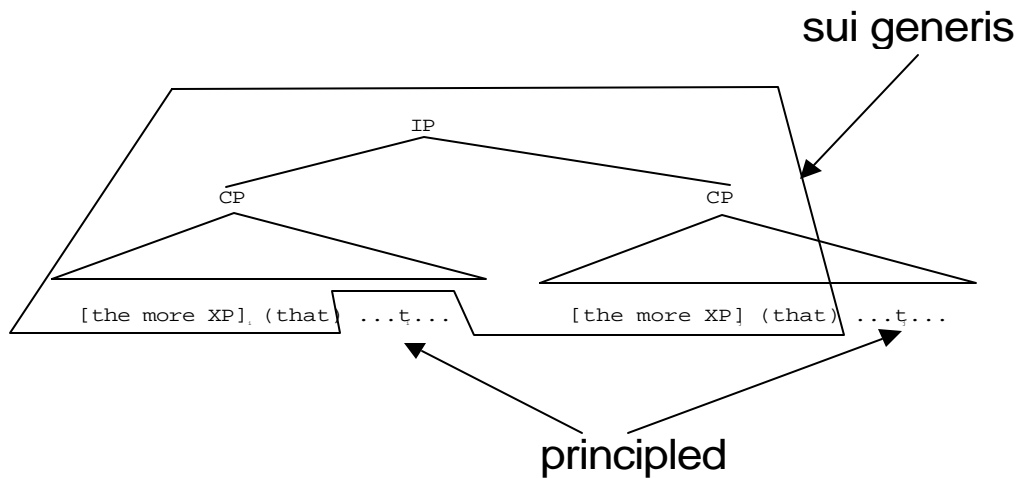
Here, there is a *sui generis* part, *the more (that) … the more (that)*, while the remainder of each clause of the construction shows all of the properties of a long distance dependency, subject to the usual constraints and principles. For example:

(10)   The more pictures she looked at t, the angrier Mary got t.

(11)   a.   The more he eats the poorer he gets.

         b.   *The more he eats [the poorer OP_i he knows a woman that gets t_i[. [CNPC, ECP]

         c.   *The more he eats [the fatter OP_i he goes to a doctor when he gets t_i]. [CED, ECP]

         d.   *The more he eats [the fatter OP_i that [that he gets t_i]] really bothers me]. [Subject Condition, ECP]

         e.   ?The more books he buys [the more books OP_i I wonder whether he'll read t_i. [Weak *wh* island]

         f.   *The more he reads [the more books OP_i I wonder to whom$_j$ he will give t_j t_i]. [Strong *wh*-island]

         g.   *The more he reads [the more people OP_i I wonder what$_j$ he will give t_i to t_j]. [Strong *wh*-island, crossing dependency]

(Some of these cases were noted by Ross 1967.)

There is nothing about the expression *the more* that would predict that it could appear in this construction. In fact, corresponding constructions in other languages use different locutions. Also idiosyncratic is the fact that *the more XP (that) …* is a CP , and that the two clauses are syntactically coordinate, points that are developed in Culicover and Jackendoff 1999. Furthermore, it is possible to say

(12)   I think that [the more she reads the smarter she gets].

which suggests that the entire expression is an IP.[1] These are properties that must simply be stipulated for the construction, in one way or the other, as part of the syntactic side of the correspondence rule.

3.3.   Idiosyncracies

In Culicover 1999b I argued that there are many idiosyncratic phenomena in natural language, at varying degrees of generality, that learners must acquire. Even if we can characterize these idiosyncrasies as the result of the interactions of very general principles, the idiosyncrasy must be marked somewhere, in order to express the fact that what occurs in English, for example *do* support, does not occur in French. I will not recapitulate this argument in detail here. However, I do want to take note of the fact that whatever the formal device for encoding this idiosyncrasy might be, the learner must extract the proper generalizations from the data as a basis for confidently configuring its grammar accordingly. This is true even if the grammar is formulated in terms of parameters; the learner must know what values to assign to the parameters, and these must be determined on the basis of the primary linguistic data.

Therefore, it seems to me that the most fundamental question, the one that comes before the question of what is the proper formalism in which to express grammatical generalizations or what is the best theory of grammar in terms of which to account for language acquisition, is to explain how it is that the learner actually extracts the relevant generalizations. Let us turn to this question.

3.4.   Extracting generalizations

I will take it as given that the task of the language learner is to arrive at an internal representation of a language such that it (the learner) is capable of producing and understanding just those expressions that the

---

[1]I am excluding the possibility of CP-recursion here.

members of the linguistic community are capable of producing and understanding. Either this representation is built into the learner to some degree, or it is extracted from the data. A strongly minimalist perspective would suggest that we assume as little as possible of one or the other. That is, either the maximal amount of knowledge is built in, or the maximal amount is acquired. Given the actual amount of variation among languages, it would seem that it would be difficult to maintain that almost everything is built in, unless of course we artificially restrict ourselves to just those phenomena that are more or less universally found. It is reasonable to suppose that the correct answer is somewhere in between, as suggested by Chomsky 1965. The problem is, where is the line drawn.

Given a particular grammatical formalism, the problem of language acquisition can be characterized as a question of whether and on what basis a language learner can arrive at a Mental Grammar that conforms to the grammatical description that we have given of it, on the assumption that there is some relationship between the two. For example, if the behavior of a Mental Grammar is best captured if we posit that it incorporates a category that contains all of the verbs, can a learner determine that such a category exists and that certain items are members of this category, or do we have to provide the learner with some prior information about this particular category and the criteria for membership in it? (In {Culicover 1999b} I argue that the answer is the former.) On the other hand, can the Language Faculty determine on the basis of experience that categories exist, or does it have to have prior knowledge that this is the case? I see no way to dispense with the assumption that it must have prior knowledge that there are syntactic categories, as well as prior knowledge of the types of evidence that bear on the category to which a word belongs.

Similar questions can be raised for every piece of knowledge, specific and general, that is embodied in the Mental Grammar. The set of things that we characterize as prior knowledge are part of the Language Faculty, or Universal Grammar. Everything else is learned. The question, now, is how?

4.   The course of language development

4.1.   Contributors to grammatical complexity

While it is not possible at this point to state definitively how language is learned, it is possible to evaluate proposals about some of the mechanisms that are at play, and the ways in which grammatical properties of the phenomena interact to influence the course of language learning. For example, a particularly complex correspondence, such as the conditional comparative discussed earlier, might plausibly be expected to be somewhat difficult to learn, in part because of the specificity of the syntactic and CS representations that are paired in the correspondence rule, and in part because of the rather sophisticated nature of the idea expressed. All of the details must be in place, and the form of the construction itself does not give any clues as to how the meaning is made up from the parts. That this is the case can be determined simply by examining the formal description of the construction in any formal framework. Of course, if the formal description is itself more opaque, e.g. the surface structure is derived from a more abstract source, such an examination is made more difficult.

Let us consider three factors that would plausibly contribute to the complexity of a sound/meaning correspondence.

C   Accessibility of the conceptual structure
C   Specificity of the syntactic structure
C   Transparency of the SS/CS mapping

These factors have an intuitive plausibility that is supported by empirical observations. For example, learners more readily acquire a lexical item that refers to a familiar concrete object than a lexical item that refers to an abstract concept. Hence we expect *cow* to be learned correctly before *idea*.[2] Prepositions denoting concrete spatial relations before those that are more abstract, and this seems to be correct (see Bowerman 1996). The greater the number of conditions on the suitability of a word to a situation, the more difficult, so we expect *cow* to be learned correctly before *assassin*. Verbal inflections that bear on the present are learned earlier than those that bear on the past, the future, or irrealis (Brown 1973). We might expect that the meaning of the comparative conditional is relatively inaccessible to a learner both from the perspective of cognitive development, and from the perspective of experience with states of affairs in which it is exemplified. Hence it is reasonable to assume as a first approximation that the accessibility of a conceptual structure is a function of its specificity, its concreteness and of the frequency of its occurrence.

Regarding the specificity of the syntactic structure, it is by no means obvious to a learner which aspects of a linguistic expression correlate with which aspects of its interpretation. Identifying which elements of a string of words and morphemes are relevant to defining a given correspondence is a non-trivial exercise. The more specific the structural description is, the greater the number of alternatives that must be excluded, all other things being equal.

Finally, let us turn to transparency. Transparency of the SS/CS mapping is related to compositionality and isomorphism. For example, suppose we have an expression of the form

(13) $Z=X^{\wedge}Y$

and the interpretation of Z is

(14) $f(Y')$

where Y' is the CS correspondence of the phrase Y. Then the correspondence in (15) can be learned by correlating X with f.

(15) $X^{\wedge}Y$
        $^{\wedge}$
        $f(Y')$

But there are many conceivable departures from this simple case.

C   It may be that some formal aspect of a string corresponds to nothing in the interpretation at all; that is, it may simply satisfy a formal condition on grammatical well-formedness. One example is verb agreement in a language like English.

C   Or it may be that there is a meaning component associated with a structure that is not associated with any of its parts, as Goldberg 1995 and Jackendoff 1997 have shown. An example would be the comparative conditional discussed earlier.

C   Or it may be that the structure of the interpretation is not isomorphic to the syntactic structure, so that a discontinuous portion of the string corresponds to some unit in the interpretation. A familiar case

---

[2]Demonstrating that a word has been correctly learned is of course far from trivial. For discussion, see Tomasello 2000 and references cited there.

is extraposition of relative clause from NP.

All of these departures from transparency of the correspondence contribute to complexity and, I would predict, to the difficulty of acquisition. It is first necessary for the learner to rule out the hypothesis of isomorphism as observationally inadequate, and then determine what the correct cross-correlations are.[3]

Note that talking in this way is neutral among grammatical theories. On any theory, it is either the case that a string X corresponds to an interpretation X', or it is not. And the correlation between substrings of X and parts of X' is similarly a question of fact. If a particular characterization of the correspondence between X and X' posits intermediate abstract representations that are not directly accessible to the learner, we can determine what computations the learner would have to go through in order to construct these representations based on the concrete evidence of sound and meaning.

4.2.   Verbal inflection and agreement

If it is on the right track, characterizing the complexity of correspondences according to their transparency suggests that there will be two types of grammatical dependencies that will be differentially accessible to learners. The one type of dependency is *functional*[4]: an aspect of the form of an element or string of elements X corresponds to and in fact depends on the role that X plays in the interpretation of the phrase or sentence. The second type of property is *formal*: the form of X does not correlate with its interpretation, and it depends only on other formal properties of the sentence in which it appears. By "form" I mean not only the phonological shape of the element itself, but its location in the linear string.

Of course this is a very gross distinction, and understanding it properly depends crucially on our being clear about what "play a role in the interpretation" means. But for certain simple examples the distinction is clear. Consider the following simple illustration of a formal dependency. In Italian, a pronominal clitic is adjoined to the right of an infinitival verb, as in (16a), and to the left of a tensed verb, as in (16b).

(16)   a.   dirlo
                 to-say-it
                 'to say it'
         b.   lo dice
                 it he-says
                 'He says it'

Moreover, the clitic can appear to the left of an infinitive in a negative imperative, such as

(17)   non lo spezzare
        not it to-break

On the other hand, in French, the pronominal clitic always precedes the verb.

---

[3]See, for example, Musolino 1998.

[4]In some cases, 'functional' is understood to pertain strictly to communicative or discourse function. I am understanding the term here as pertaining to the interpretation associated with a form.

(18)   a.   le dire
              it to-say
              'to say it'
        b.   il le dit
              he it says
              'He says it.'

In Italian but not in French there is a formal dependency between the position of the clitic and the type of verb. Crucially, in neither case is there a functional dependency, since the interpretation of the clitic (e.g. as a direct object) in neither language depends on its position in the sentence or the type of verb that it is adjoined to.

Along similar lines, the form of an irregular verb is an arbitrary fact that must be learned, and hence is strictly formal and not functional. It does not correspond to the interpretation of the verb except insofar as it signals that the verb falls into a certain position in a paradigm. The intuition that I am working with is that the functional correspondence of form with with meaning is a more immediate clue to the learner about how the syntax works than formal correlations between elements in a syntactic structure. Overgeneralizations of regular morphology are in fact widely attested in the acquisition of a wide range of formal phenomena. But it is plausible that these overgeneralizations occur precisely because there is no functional basis to the distinction. That is, there is no meaning difference associated with the formal difference; the learner is not immediately aware that the form matters, or what the precise conditions are that determine the form. In such cases, we would expect that the statistically most dominant alternative wins, if there is one; otherwise there will be various strategies adopted by learners.[5]

Another example of the functional/formal distinction concerns finite subject-verb agreement in English and Spanish. Here we can see that there is more going on than just frequency of occurrence. In English the form of the finite verb is predictable from the properties of the subject with which it agrees. Errors in agreement produce sharp ungrammaticality, but do not affect interpretation at all.

(19)   John /is/*are/*am/ here.

The same is true in Spanish when there is an overt subject.

(20)   Esta/*estan/*estoy   Juan aqui.
          is/are/am   Juan here
        'John is here.'

However, in Spanish a pronominal subject is not always overtly expressed; there is 'prodrop.' In this case, the overt agreement morphology conveys crucial semantic information about the reference of the subject.

(21)   a.   Esta aqui
              he-is here
              `He is here.'
        b.   Estan aqui
              they-are here

_____

[5]See Stromswold 1990.

10

'They are here.'
c.   Estoy aqui
      I-am here
      'I am here.'

Here, Spanish verbal morphology is functional. English verbal morphology, on the other hand, is formal.

Other languages that fall into the same class as English are German, Dutch, and French. In these languages, the form of a tensed verb depends on the person and number of the subject of the sentence; an infinitival verb bears a different form. In languages such as Swedish and Danish verbal agreement has been lost entirely; however, the form of the verb does depend on whether it is finite or nonfinite. The data can be roughly summarized as follows.

(22)   1.   *English*
              Finite
                        Irregular verbs: marked agreement
                        Regular verbs: base form throughout except for marked 3rd person singular
                                  agreement
              Nonfinite: base form
        2.   *German/Dutch/French*
              Finite: marked agreement
              Nonfinite: no agreement, infinitival form
        3.   *Swedish/Danish*
              Finite: no marked agreement - base form
              Nonfinite: no agreement, infinitival form

Contrast these languages with those that have prodrop.

        4.   *Spanish/Italian*
              Finite: marked agreement, prodrop
              Infinitive: no agreement, infinitival form

In prodrop languages, since verbal inflection is functional, it should be learned relatively quickly, if not instantly, since the verbal inflection plays a functional role in expressing predication. The evidence in the literature that bears on this question suggests that in the prodrop languages, verbal morphology is learned immediately and without significant error (see, for example, Hamann, et al.1996.) As soon as predication develops, the learner recognizes that the inflection on a verb correlates with the missing subject in cases of prodrop. The same correlation holds when there is an overt subject. We therefore predict quite specifically that there will be few if any cases in prodrop languages of nonfinite verbs used in main clauses with overt subjects, or of finite verbs used where there can only be a nonfinite verb. This prediction is confirmed by the observation of Hoekstra and Hyams 1995 that the occurrence of root infinitives, that is, nonfinite forms in positions where the language requires finite forms, is in fact very rare in the prodrop languages; see also Wexler 1998. The data are summarized in Table 1, taken from Hoekstra and Hyams 1995.

	Child	Age	%RIs
Languages			
Italian	Diana	2;0	.0
(Guasti 1992)	Martina	1;11	.16
		2;1	.04
Italian	Paoa	2;0-2;5	.07
(Schaeffer 1990)	Daniele	1;7-2;6	.08
	Massimo	1;7-2;6	.06
	Gabriele	1;7-2;6	.07
	Orietta	1;7-2;6	.05
	Elisabet	1;7-2;5	.10
	Frances	1;9-2;5	.05
Spanish	Damariz	2;6-2;8	.05
(Grinstead 1993)	Juan	1;7-2;0	.12
		2;1-2;4	.10
Catalan	Cuillem	1;11-2;6	.03
(Torrens 1992)	Marti	2;0-2;5	.03

Table **1**. Root infinitives in prodrop languages, from Hoekstra and Hyams 1995.

On the other hand, the distinction between finite and nonfinite forms in German, Dutch and French a formal one; the data in Table 2 from Hoekstra and Hyams (1995) show that the incidence of root infinitives in place of finite verbs is much higher in these languages.

	Child	Age	%RIs
French	Nathalie	1;7-2;11	.76
(Pierce 1992)	Daniel	1;5-2;5	.60
English	Eve	1;6-1;10	.78
Swedish	Freja	1;11-2;0	.38
(from Guasti 1992, based on Platzack)	Tor	1;11-2;2	.56
German	S	2;1	.46
(from Guasti 1992 based on Weissenborn)		2;2	.4
Dutch	Laura	1;8-2;1	.36
(Weverink 1989)	Tobias	1;10-1;11	.36
	Fedra	1;10-2;1	.26

Table **2**. Root infinitives in non-prodrop languages, from Hoekstra and Hyams 1995

4.3. English Subject-Aux Inversion (SAI)

We can see the role of the formal/functional distinction even within a single language. Consider the problem of learning Subject-Auxiliary Inversion in English. As the following sentences show, in main clause yes-no questions and *wh*-questions both, the auxiliary verb appears before the subject.

(23) a. What is your name?
      b. Are you my friend?
(24) a. *What your name is.
      b. (*)You are my friend.

At a certain stage of language development, children learning English systematically produce sentences of the form (24a) while at the same time they are producing sentences of the form (23b). Klima and Bellugi 1966 report that this pattern of behavior is found in their Period 2, which corresponds to the following ages.

(25) *Period 2 (MLU 2.25-2.75)*
      Adam   2;7 - 2;11
      Eve   1;9 - 1;11
      Sarah   2;8 - 3;1

At the end of this period, inversion begins to appear in *wh*-questions as well as in yes-no questions.

Note that inversion in *wh*-questions is formal; it is simply a formal property of *wh*-questions that the auxiliary must appear in second position in English. On the other hand, inversion in yes-no questions is functional, in the sense that it distinguishes a yes-no question from the corresponding declarative.[6] It follows from the assumption that learning of formal dependencies is slower than the learning of functional dependencies that inversion will emerge later in *wh*-questions than in yes-no questions in English speaking children.[7]

It is important to stress here the contrast between this type of account and the type of account that is typical in parameter setting theory. In parameter setting theory, the formal characterization of inversion in the two types of questions is identical, at least to a first approximation. In current terms, a functional head, e.g. Infl, with the feature [+F] adjoins to the head C so that C can discharge the feature f+F$_n$. To the extent that it takes a certain amount of time for the learner to acquire the knowledge that C has to discharge the feature f+F$_n$ and that in a question, Infl has the feature [+F], there is no particular basis for predicting that *wh*-questions and yes-no questions will be distinguished. It is of course possible to stipulate that there is a difference, but such a stipulation lacks explanatory force. For example, there could be two features, f+F$_n$ and f+G$_n$, one for *wh*-questions and one for yes-no questions. This is simply a way to describe the problem, and does not really solve it.

Alternatively, we could explore an alternative formulation of the syntax of these constructions and stipulate that the early syntactic structure in some way rules out the simultaneous application of *wh*-Movement and SAI (e.g. they could both involve movement to the same landing site in the early grammar, and subsequently a new landing site is acquired for both the wh and the Infl[8]). Such an account would also be stipulative in its assumptions about the form of early grammar and the basis for the transition to an adult grammar.

In fact, it seems to me that the simplest hypothesis is that the learner is capable of dealing with the syntax of inversion at the point at which yes-no questions have been learned, but that inversion has a different functional status in the two types of questions that accounts for the fact that it emerges later in the *wh*-questions than in the yes-no questions.

4.4.   Referential and expletive pronouns

Let us consider one more example, the difference between referential and expletive pronouns. A referential pronoun has a correlate in conceptual structure, while an expletive pronoun simply satisfies a formal grammatical requirement. We would predict, then, that in languages with both, referential pronouns would be used earlier than expletive pronouns.[9]

---

[6]Another limited way of marking a sentence as interrogative is through intonation.

[7]I am simplifying somewhat. Given that learners are conservative, inversion should develop in particular types of *wh*-questions and gradually generalizes. A learner may produce inversion in a *wh*-question only with a particular *wh*-form and verb, e.g. *what did NP VP*. See Kuczaj and Brannick 1979 and Stromswold 1990 for evidence that this is in fact the case.

[8]See, for example, Klein 1982.

[9]Thanks to Christina Schmitt (p.c.) for suggesting this distinction.

A preliminary examination of the child language data in the CHILDES database[10] suggests that this prediction is on the right track. Brown and his colleagues studied three children, Adam, Eve and Sarah. The pronoun *it* appears in the early transcripts of each of the children (in the first of Adam [2;3.4] and Eve [1;6.0], and the third of Sarah [2;3.19]). In the transcripts of Adam, there are just two instances of the expletive *it*, one at 4;5.0, and one at 4;10.0, as shown in table **3**. When he uses a predicate that takes expletive *it* in adult English, Adam drops the expletive, as illustrated in table **4**. On the other hand, we find uses of expletive *it* in the caretaker's speech as early as the first recorded interaction (where Adam is 2;3.4) and throughout. Some examples are given in Table **5**.

```
@Comment:       ---------------------------------------
@Age of ADA:   4;5.0
@Comment:       *** File "Adam48.cha": line 221. Keyword: it
*ADA:  it looks like I'm dere .

@Comment:       ---------------------------------------
@Age of ADA:   4;10.0
@Comment:       *** File "Adam54.cha": line 1169. Keyword: it
```

Table 3. Earliest occurrences of expletive *it* in Adam's speech

```
@Comment:       *** File "Adam01.cha":
*MOT:  is it sunny outside ?
*ADA:  no sunny outside .
```
Table 4.   Example of Adam's omission of expletive *it*

---

[10]MacWhinney (1995).

```
@Comment:          ---------------------------------------
@Comment:          *** File "Adam01.cha": line 1196. Keyword: it
*MOT:  is it sunny outside ?

@Comment:          ---------------------------------------
@Comment:          *** File "Adam03.cha": line 373. Keyword: it
*MOT:  is it raining ?
@Comment:          ---------------------------------------
@Comment:          *** File "Adam03.cha": line 386. Keyword: it
*MOT:  is it cloudy ?
@Comment:          ---------------------------------------
@Comment:          *** File "Adam03.cha": line 389. Keyword: it
*MOT:  what kind of day is it ?

@Comment:          ---------------------------------------
@Comment:          *** File "Adam05.cha": line 1447. Keyword: it
*MOT:  it was raining .

@Comment:          ---------------------------------------
@Comment:          *** File "Adam20.cha": line 489. Keyword: it
*MOT:  well # it doesn't seem to have a top .

@Comment:          ---------------------------------------
@Comment:          *** File "Adam20.cha": line 1354. Keyword: it
*MOT:  it seems to me that you tire of your games very easily .

@Comment:          ---------------------------------------
@Comment:          *** File "Adam34.cha": line 1150. Keyword: it
*MOT:  no # it seems that something must have been broken .

@Comment:          ---------------------------------------
@Comment:          *** File "Adam53.cha": line 705. Keyword: it
*MOT:  yes # it seems a little too quick .
```

Table 5.   Examples of expletive *it* in caretaker's speech to Adam

The word *there* in English is also an expletive, but one that has a conceptual structure function, that of expressing existence. Therefore we would expect this expletive to appear earlier in Adam's speech, and it does, at 2;3.4. (See Table **6**.)

```
@Comment:          ---------------------------------------
@Age of ADA:  2;3.4
@Comment:          *** File "Adam01.cha": line 251. Keyword: there
*ADA:  there go one .
```

Table **6**. Earliest occurrence of expletive *there* in Adam's speech

With the exception of three instances of

(26)   *EVE: what time it is

there are no expletive pronouns in the Eve transcripts, and there are none in the Sarah transcripts. The dialogue in Table 7 suggests that Eve's use of expletive *it* may be due to the repetition of a fixed formula.

```
*RIC:     what's that?
*RIC:     a watch.
*RIC:     it says what time it is.
@Tape Location: 715
*EVE:     where xxx yours?
%act:     moving over to Colin
*COL:     see?
*EVE:     what time it is.
*RIC:     hmm?
*EVE:     what time it is.
%spa:     $RES
*RIC:     uhhuh.
*RIC:     it tells what time it is.
*EVE:     what time it is.
%spa:     $RES
*RIC:     uhhuh.
*RIC:     it tells what time it is.
*EVE:     Fraser # I go have lunch I go have lunch.
```

Table 7. Dialogue with Eve about "what time it is"

Expletive *there* first appears when Eve is 1;6.0 and when Sarah is 3;2.10. The first instance of expletive *it* in Bloom's transcripts of Peter is at 2;6.16, while the first instance of expletive *there* is possibly at 1;10.0 --

(27)   *CHI: there it is .
         [File: peter01.cha]

–- and definitely at 2;1.0.

(28)   *CHI: there's a tape right there # huh.
         [File: peter07.cha]

The occurrences of the various forms for these children is summarized in table 8.

17

	First ref. pronominal *it*	First expletive *it*	First *there*
Adam	2;3.4	4;5.0	2;3.4
Eve	1;6.0	???	1;6.0
Sarah	2;3.19	???	3;2.10
Peter	1;9.8	2;6.16	1;10.0

Table **8**. Occurrence of pronouns and expletives in transcripts of some English speaking children

As far as I can determine, the same pattern holds in other languages with expletive pronouns. So, in Clahsen's German transcripts I find only two instances of expletive *es*, as given in Table **9**. In Wagner's German transcripts we find the examples in Table **10**. The first instance of French expletive *il* in Champaud's transcript of Grégoire is that shown in Table **11**. *Il y a* appears at approximately the same age. The expletive *il y a* appears in Philippe's language first at age 2;2.26, while expletive *il* used otherwise first appears no earlier than 2;8.29 in the expression *il fait rien*.

```
From file <Mat21.cha>
@Age of MAT: 3;0.21
*MAT:  es gibt ein turm .
From file <Mat22.cha>
@Age of MAT: 3;1.21
*MAT:  da gibt ein turm .
```

Table **9**. First German expletives, Mat.

```
From file <Andreas2.cha>
@Age of AND:    2;1.
*AND:  ## es kommt kein Ball, Mama .
```

Table **10**. First German expletives, Andreas

```
From file <Greg09.cha>
@Age of GRE:    2;5.13
*GRE:    il fait froid sur ça

From file <Greg10.cha>
@Age of GRE:    2;5;27
*GRE:    euh, euh, il y en a deux .
```

Table **11**. First expletives, Grégoire

While the absolute age at which the expletives appear seems to vary dramatically from one learner to the next, the pure expletive corresponding to the English *it* appears later than the existential expletive, which in turn appears later than the referential pronoun. So, to the extent that there is evidence (and of course much

more research needs to be done), the predicted sequence is the one that occurs.

5.   **Conclusions**

To conclude, I have argued for the following points.

C   The right way to think of grammar is that it is an idealized description of the behavior of the device that embodies knowledge of language.
C   Knowledge of language consists of sound/meaning correspondences.
C   Many different formal ways of expressing linguistic knowledge can be equivalent in terms of the sound/meaning correspondences that they describe.
C   Beyond this, a grammatical description should characterize the sound/meaning correspondence in such a way that it correctly fixes its complexity with respect to other correspondences. A correlate of this complexity is the ease with which it is acquired by language learners on the basis of their linguistic

I have also suggested a Dynamical perspective, in which we distinguish between the architecture of the Mental Grammar in the mind and the description of the behavior of this device. To a considerable extent this view divorces linguistic theory and the theory of language acquisition. The device that is actually in the mind must produce expressions, and structures, and generalizations that are consistent with the description permitted by the linguistic theory, but the theory does not constitute a strong claim about the architecture of what is in the mind. In particular, on this view the course of language acquisition is accounted for by the architecture of what is in the mind, not by the architecture of the grammatical theory. Of course, to the extent that the linguistic theory reflects the architecture of the Mental Grammar, there will be a correspondence between the two, but it is not a principled one.

References

Bowerman, M. (1996) "Learning How to Structure Space for Language: A Crosslinguistic Perspective," in Bloom, P., M. A. Peterson, L. Nadel and M. F. Garrett, eds., *Language and Space*, MIT Press, Cambridge, MA.

Brown, R. S. (1973) *A First Language*, Harvard University Press, Cambridge, MA.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.

------ (1975) *The Logical Structure of Linguistic Theory*, Plenum, New York.

Culicover, P. W. (1999a) "Minimalist Architectures: Review Article of *The Architecture of the Language Faculty*, by Ray S. Jackendoff," *Journal of Linguistics* 35, 137-150.

Culicover, P. W. (1999b) *Syntactic Nuts*, Oxford University Press, Oxford.

------ and R. Jackendoff (1999) "The View from the Periphery: The English Comparative Correlative," *Linguistic Inquiry* 30, 543-571.

Goldberg, A. E. (1995) *Constructions: A Construction Grammar Approach to Argument Structure*, University of Chicago Press, Chicago.

Hamann, C., U. Frauenfelder and L. Rizzi (1996) "Evidence from the Augustin Corpus." in Koster, C. and F. Wijnen, eds., *Proceedings of GALA 1995*, Center for Language and Cognition, Groningen.

Hoekstra, T. and N. Hyams (1996) "Missing Heads in Child Language," in Koster, C. and F. Wijnen, eds., *Proceedings of GALA 1995*, Center for Language and Cognition, Groningen, 251-260.

Jackendoff, R.(1983) *Semantics and Cognition*, MIT Press, Cambridge.

------ (1997) *The Architecture of the Language Faculty*, MIT Press, Cambridge, MA.

Klein, S. M. (1982) *Syntactic Theory and the Developing Grammar*, unpublished doctoral dissertation, UCLA, Los Angeles.

Klima, E. and U. Bellugi (1966) "Syntactic Regularities in the Speech of Children," in Lyons, J. and R. Wales, eds., *Psycholinguistic Papers.*, University of Edinburgh Press, Edinburgh, 183-206.

Kuczaj, S. and N. Brannick (1979) "Children's Use of the Wh Question Modal Auxiliary Placement Rule," *Journal of Experimental Child Psychology* 28, 43-67.

MacWhinney, B. (1995) *The CHILDES Project : Tools for Analyzing Talk*, L. Erlbaum, Hillsdale, N.J.

Musolino, J. (1998) "Universal Grammar and the Acquisition of Semantic Knowledge: An Experimental Investigation Into the Acquisition of Quantifier-Negation Interaction in English," Unpublished doctoral dissertation, University of Maryland, College Park.

Ross, J. R. (1967) *Constraints on Variables in Syntax*, unpublished doctoral dissertation, MIT, Cambridge, MA.

Stromswold, K. (1990) "Learnability and the Acquisition of Auxiliaries," unpublished doctoral dissertation, MIT, Cambridge, MA.

Tomasello, M. J. (2000) "Do Young Children Have Adult Syntax," *Cognition* 74, 209-253.

Wexler, K. (1998) "Very Early Parameter Setting and the Unique Checking Constraint: A New," *Lingua* 106, 23-79.