6 Competition, Variation, and Language Learning

Elizabeth Bates
University of California, San Diego

Brian MacWhinney
Carnegie-Mellon University

The problem of accounting for the acquisition of language can be decomposed into two smaller problems: how to account for what is universal in language development and how to account for what is variable. We have seen a number of elegant and detailed accounts of universal processes in language acquisition. But these models have not yet taken seriously the existence of two significant aspects of variation in the acquisition process: *1.* variation across natural languages and *2.* variation between individual learners within a particular language. In this paper we discuss a model of language acquisition that has attempted to deal with the first type of variation. We then indicate how this model will have to be elaborated in order to deal with the second form of variation, i.e. variation between individual learners. The development of a detailed mechanistic account for variation is particularly important for those who are interested in a biologically-based theory of language learning. By looking at variation, we are addressing a fundamental issue in the biological sciences, the plasticity of developing systems. How many different forms can a biological system take under normal and abnormal conditions?

The data currently available (Slobin, 1985) provide little evidence for a single, universal sequence in the acquisition of basic grammatical forms. This means that, in order to construct a universalist account of the acquisition of grammar, one must introduce concepts that extract universal patterns out of what appear to be particularistic data. One way of doing this is to think in terms of the nativist concept of "parameter-setting" (e.g. Chomsky, 1982; Lightfoot, 1982). According to a parameter-setting analysis, natural languages vary too much in their basic structure to permit a definition of universals entirely in terms of some *intersect,* i.e. the set of structures that *every* language has to have. Rather than obeying categorical universals, languages are governed by "implicational uni-

versas," a pool of structural possiblities in which any choice carries important structural consequences of the "If X, then Y" variety. Each individual language has charted a path through this set of possibilities. However, given the many implicational constraints within the system, the total set of possible pathways is finite, and rather small. According to this view, language acquisition can be viewed as a process of the successive setting of parameters in a way that allows the system to live with the preordained consequences of each setting. Biology provides the universal parameters; language input triggers a set of constrained choices within that pool of possibilities.

The problem with parameter setting as a model of cross-language variation is that it predicts sudden and all-or-none decisions, carried out in a single specified order, with essentially no opportunity to turn back once a parameter is set. Furthermore, the model is based on the assumption that the adult "steady state" can be modelled in terms of the presence or absence of certain structural types. We will present cross-linguistic evidence to suggest that languages vary not only in their end points (as a parameter-setting theory would predict) but also in the initial hypotheses that children hold about their grammar. This evidence indicates that the sequence of "parameter testing" is apparently not universal. Furthermore, the passage from initial states to end states is a gradual one. Two competing tendencies may coexist for prolonged periods of time, cycling in and out as though the child were unable to make up her mind. Finally, the "steady state" reached by adults also contains patterns of statistical variation in the use of grammatical structures that cannot be captured by discrete rules. This kind of cross-linguistic variation is difficult to capture with an all-or-none model.

Recognizing many of the problems we have just noted, Pinker (this volume) attempts to modify the parameter-setting model by encoding the parameters themselves in probabilistic terms. But this approach underestimates the depth of the problem. Any model that rests exclusively on universalist principles will fail to provide veridical accounts of the variable facts of language acquisition. Any model that rests solely on variation would also fail to provide coherent explanations for universal patterns. What is needed, instead, are models that are fundamentally capable of expressing both the universal and the variable aspects of language acquisition.

We have called the model that we will be presenting the Competition Model. Like other data-driven, connectionist models, the Competition Model allows statistical properties of the input to play a major role in determining order of acquisition as well as the nature of the final state. In this way the fundamental mechanisms of the model provide us with a way of understanding variability in particular words, segments, or constructions. This emphasis on ways in which the organism can adapt to the shape of the input allows us to apply our model to the study of cross-linguistic variation.

The other major type of variation that we will consider is variation in the contour of the learning process. Here, the major source of variability is the child

herself. It has become increasingly clear in recent years that children can acquire English in radically different ways—at least in the early stages of language learning (Nelson, 1981; Bretherton, McNew, Snyder & Bates, 1983; Bates, Bretherton & Snyder, 1985). Furthermore, these differences are apparently not due to variation in the child's linguistic input, although environmental factors can serve to discourage or enhance a particular linguistic "style" once it becomes apparent (Nelson, 1973; Furrow & Nelson, 1984; Goldfield & Snow, in press). Bates et al. (1985) show that individual differences in language development can be brought about by the differential strength and/or differential timing of two or more underlying mechanisms responsible for language acquisition and language processing. We are thinking of dissociable processes as faculties in the sense of abilities or skills. For example, one child may have a well-developed faculty for memorizing strings with detailed phonology, whereas another child may make greater use of a faculty for analyzing these same strings into their component parts. Differences in the patterns of development between these two children can then be seen as a reflection of differential use of these basic faculties. One major goal of our work with the Competition Model is to delimit a set of fundamental processing mechanisms whose strength at a given point in development varies across learners in a way that can eventually be linked to fundamental differences between the learners.

In the first part of our paper, we will consider the treatment of data on cross-linguistic variation in language learning provided by the Competition Model. In the second part of the paper, we will consider evidence for individual variation in the acquisition of English. We will then consider some suggested modifications that could be added to the competition model or any other connectionist/lexicalist theory to account for these patterns of variation.

VARIATION ACROSS LANGUAGES

Here we will examine the problem of accounting for the details of the differences in the course of language acquisition between children learning different languages. We will provide a sketch of the Competition Model, with emphasis on the principle of cue validity and the predictions that it makes for cross-linguistic differences in language learning. Then we will present the cross-linguistic evidence for and against a simplistic version of the model leading to the postulation of two kinds of developmental constraints on cue validity: functional readiness and cue cost.

The Competition Model

The competition model is a particular instantiation of a general functionalist approach to language performance and language acquisition. As defined by

MacWhinney, Bates, and Kliegl (1984, p. 128), functionalism is the belief that "The forms of natural languages are created, governed, constrained, acquired and used in the service of communicative functions."

From this point of view functionalism is the natural alternative to theories that postulate a severe separation between form and function in the grammars of natural languages. The idea that grammars routinely and generally spawn and proliferate forms that play no role in facilitating communication is foreign to the functionalist position. While recognizing that some systems, such as the gender-case marking of German article declension, may have lost much of their original function, we believe that language works continually to find functions for forms that have lost their original use. There are, of course, many versions of the functionalist approach with different kinds of claims requiring different kinds of evidence:

- at the diachronic level where functions play a role in the evolution of a particular language,
- at the synchronic level where functions continue to constrain linguistic forms in real time comprehension and production,
- at the developmental level where children use communicative functions as a guide in the acquisition of forms,
- at the level of formal grammar where rules in the grammar make direct reference to semantic and pragmatic symbols.

The competition model makes functionalist claims at the first three of these four levels. In other words, it is not offered as a formal model of **linguistic competence** but rather as a model of **linguistic performance.** This concentration on performance has one particularly important implication: in modelling the differences among natural languages, our goal is to provide an explicit account not only for the kinds of discrete "yes or no" phenomena that play a role in traditional linguistic models but also for the probabilistic differences between natural languages that are observed in real-time language use. In other words, **we are focussing on cross-linguistic variation in the mapping between form and function in language comprehension, production and acquisition.**

Before we describe the current version of the competition model, let us first consider some illustrative contrasts between two of the languages that we have studied in greatest detail, Italian and English. We will concentrate on a small but very important aspect of the grammar: the structural phenomena associated with the form of "sentence subject" as they relate to basic functions such as agent/actor and patient/object.

In both Italian and English, the "basic" or pragmatically-neutral word order is Subject-Verb-Object (SVO). There are no case inflections to mark semantic relations, except for some remnants of case in the pronoun system (e.g., the

contrast between *I* and *me*). These are both Indo-European languages, and they share a large number of cognates and word-formation patterns. But despite such formal similarities, Italian and English behave quite differently in everyday use.

First of all, Italian permits a great deal of pragmatic variation in basic constituent order. In fact, every logical order of subject, verb and object can be found in informal speech. This is illustrated in the following excerpt from Bates, MacWhinney and Smith (1983):

1. SVO: lo mangerei un primo. (I would eat a first course.)
2. OSV: La pastasciutte Franco la prende sempre qui. (Pasta Franco it orders always here.)
3. VSO: Allora, mangio anche io la pastasciutte. (Well then, am eating also I pasta.)
4. VOS: Ha consigliato la lasagna qui Franco, no? (Has recommended the lasagna here Franco, no?)
5. OVS: No, la lasagna l'ha consigliata Elizabeth. (No, the lasagna it has recommended Elizabeth.)
6. SOV: Allora, io gli spaghetti prendo. (In that case, I the spaghetti am having.)

The flexibility of word order is compounded still further by the fact that Italian is a "pro-drop" language, i.e., a language in which subject omission is a perfectly legal and very common option (occurring in approximately 70% of the clauses in informal speech among adults, according to Bates, 1976). As a result, the most frequent form in Italian discourse is not SVO, but (S)VO or O(S)V. Given this combination of word order variation plus ellipsis, the identity of subject and object is not at all predictable in Italian by word order information alone.

How do Italians get away with such behavior? For one thing, they can often rely on a richly marked system of verb morphology to let them know "who did what to whom." This contrasts markedly with the degraded system of verb morphology in English, as illustrated below:

1. Io mangio. I eat.
2. Tu mangi. You-informal-singular eat.
3. Lui/Lei mangia. He/She/You-informal-singular eat.
4. Noi mangiamo. We eat.
5. Voi mangiate. You-informal-plural eat.
6. Loro mangiano. They/You-formal-plural eat.

In these present tense examples, the only available contrast in English is provided by the third person singular "'s." The other forms are entirely ambiguous.

Although ambiguities do occur in Italian (e.g., between third person and the second person formal), they are relatively infrequent. In a sense, then, Italians can make a "case-like" use of verb agreement and other aspects of morphology (e.g., the system of pronominal object clitics that agree with the object in person, gender and number) as primary cues to semantic role relations. Quite simply, Italians can "trust" morphology more than they can "trust" word order—precisely the opposite of the patterns expected and observed in English.

There are some other contrasts between English and Italian in the way the subject system works, contrasts that are particularly hard to capture in terms of the presence or absence of structures and/or rules. For example, Italians seem to have an aversion for constructions with an indefinite subject, even though such constructions are perfectly "grammatical" in the strict sense (Devescovi and Taeschner, 1985). To avoid indefinite subjects, they often make use of complex, syntactic structures that an English speaker prefers to avoid. For example, given a picture of a monkey eating a banana, English speakers from 2 to 90 years of age are likely to say,

A monkey is eating a banana.

Given exactly the same picture, Italians as young as 2 years of age are six to seven times more likely to say something like,

There is a monkey that is eating a banana.

These facts are really not mysterious when we remember that Italians are making extensive use of word order variation as a pragmatic device. In the terminology of functional grammar, we would say that Italian is a relatively *topic-dominant* language (Li & Thompson, 1976; Schachter, 1976). The subject system is associated more strongly with the functional notion of discourse topic, and hence it is more likely to be used to talk about given or established information. Insofar as an indefinite determiner is used to mark new information, indefinite subjects represent a violation of "topichood." In the Philippine languages described by Schachter, indefinite subjects are not only avoided, they are entirely ungrammatical. To convey the information contained in our monkey/banana picture, these Philippine speakers would be required to begin by introducing the monkey in some kind of existential clause, adding the verb comment only after the subject is established in discourse (e.g., "There is a monkey, and he is eating a banana," or "I see a monkey that is eating a banana"). From this point of view, the Italian hatred for indefinite subjects and love of relative clauses suggest that Italian represents a statistical midpoint between the subject system of English and the subject system of Philippine languages.

We can use these performance facts about English and Italian to "walk through" the main features of the competition model as we currently understand it. The model can be summarized in terms of the following list of claims.

Direct Mapping

Only two levels of processing are specified in this performance model: a *functional* level (where all the meanings and intentions to be expressed in an utterance are represented) and a *formal* level (where all the surface forms or expressive devices available in the language are represented). Mappings between these levels are said to be direct. The notion of language as a system of mappings between form and function is the basic insight underlying the linguistics of de Saussure. In this sense, much of our work can be seen as an application of the Saussurian framework to the domain of performance or *parole*.

The principle of direct mapping does *not* require that the relationships between form and function stand in a one-to-one relation. Rather, direct mapping means that it is possible for languages to integrate on a single level cues that refer to different data types. In sentence comprehension, the parser is able to consider compounds or configurations of lexical semantic cues (e.g., animacy), morphological cues (e.g., agreement markers), word order cues (e.g., preverbal position), and intonational cues (e.g., contrastive stress). As we will see later, the parser may also take into account low-level acoustic/phonetic information about the "perceivability" and hence the "trustworthiness" of particular markers. In sentence production, the same configurations can be retrieved and assembled together, as a block. This contrasts with modular theories in which each distinct data type is handled by a separate processor (Fodor, 1983; Garrett, 1981).

In this paper, we will talk about the elements at each level in essentially "pretheoretical" terms, using global notions like "topic" and "agent" to describe the level of function and theory-neutral terms like "subject-verb agreement" and "contrastive stress" to describe the level of form. MacWhinney (this volume) will provide a more formal and explicit characterization of these two levels in terms of a system of lexical items (to describe the level of form) and a propositional notation (to describe the level of function). At this point, however, it is probably worth pointing out that the principle of direct mapping is basically a *lexicalist claim*. Like Saussure, we focus on the relation between the sign and the signified as the central structure controlling language processing. Extending Saussure's notion of lexically-based syntagmatic relations, we show how the native speaker learns to map phrasal configurations onto propositions, using the same learning principles and representational mechanisms needed to map single words onto their meanings.

Cue Validity

The major predictive construct in the competition model is cue validity. Following Brunswik (1956), we argue that human beings possess psychological mechanisms that bring them in tune with the validity or information value of cues in their ecology. This means that validity is an objective property of the cue

itself, i.e., a property of the perceptual environment relative to some organismic state.

MacWhinney (1978) and MacWhinney, Pleh, and Bates (1985) have analyzed cue validity into two components: **cue availability** (i.e., how often is this piece of information offered during a decision making process?), and **cue reliability** (i.e., how often does the cue lead to a correct conclusion when it is used?). McDonald (1984) has tested some alternative methods for calculating cue validity.

1. *Availability* in McDonald's scheme is best expressed numerically as the ratio of the cases in which the cue is available over the total number of cases in a given task domain. For example, the availability of preverbal position is very high in English but relatively low in Italian. This reflects the fact that subjects are frequently omitted in Italian leaving many verbs in sentence-initial position.
2. *Reliability* can be expressed numerically as a ratio of the cases in which a cue leads to the correct conclusion, over the number of cases in which it is available. For example, preverbal position is a highly reliable cue in English where it is almost always assigned to the agent of a transitive action; it is a very unreliable cue in Italian (when it is available at all), since OV and SOV constructions are both possible and likely.
3. *Validity* is defined as the product of availability times reliability. Given the reliability and availability calculations described above, this necessarily means that the cue validity of preverbal position is very high in English and very low in Italian—a fact that is reflected in the performance of English and Italian listeners in the experiments described below. Insofar as validity is a property of the environment relative to some organismic state or goal (e.g., the communicative function of "agent"), we can calculate validity directly from samples of the linguistic input to children.

Cue Strength

To model the organism's knowledge about the validity of information, we postulate a subjective property of the organism called **cue strength.** This is a quintessentially connectionist notion, referring to the probability or weight that the organism attaches to a given piece of information relative to some goal.

In our psycholinguistic instantiation of this idea, each link between a given surface form (e.g., a nominative case marking) and an underlying function (e.g., the agent role) is given a weight or strength. With this kind of mechanism no sharp line is drawn between probabilistic tendencies and deterministic rules. An obligatory relationship between form and function is nothing other than a connection whose strength approaches unity. This permits us to capture statistical differences between adult speakers of different languages, e.g., the tendency for

English listeners to "trust" word order more than their Italian counterparts. It also permits us to capture facts about language change (in language history and/or in language learning) in gradual and probabilistic terms; we are not forced to postulate a series of all-or-none decisions, i.e., moments where parameters are definitively set and rules are added or dropped.

However, cue strength is not completely isomorphic with cue validity. Calculations of cue strength require some specification of the size and frequency of the task domain. For example, in our research we have calculated cue validity from one language to another within the domain of sentences that require a decision about who did what to whom, i.e., the domain of transitive sentences. Some tasks—like this one—are very frequent; others—like some of the tasks used in Piagetian studies of scientific reasoning—rarely come up. Cue strength will be a function of *both* cue validity and task frequency. In practical terms, it is of course quite difficult to estimate the frequency of a task with any precision. Sometimes, the best that we can do is to offer an *ordinal* prediction that cues for highly infrequent tasks will be learned relatively late as compared with cues for frequent tasks. However, within a specified task domain, the major determinant of order of acquisition and eventual cue strength should be cue validity. This leads to a series of very strong and falsifiable claims about language learning.

Horizontal and Vertical Correlations

We do not believe that mappings can only be between forms and functions. Rather, in the Saussurian system of forms and functions, three types of correlations are possible.

1. There may be direct *vertical correlations* between forms and functions. For example, the form of preverbal positioning in English is correlated with the function of expressing the actor role.
2. There may be correlations between forms themselves. These are *horizontal correlations* on the level of form.
3. There may be correlations between functions. These are *horizontal correlations* on the level of function.

In Bates and MacWhinney (1982) we focus our attention on the role of vertical correlations in language learning. However, we also recognize the importance of horizontal correlations. We speak of horizontal correlations on the level of function in terms of systems of competition between functions governed by "divide the spoils" and "peaceful coexistence" solutions. We speak of horizontal correlations on the level of form in terms of ossified forms. Maratsos (1982, 1983) is correct in pointing out that the relegation of form-form correlation detection to the bone pile of language history is an overly strong application of functionalism. We have been persuaded by his emphasis on the detection of correlation as a

fundamental process in language acquisition. However, we also believe that the child does not consider *all* possible correlations between *all* items in *all* sentences in acquiring an accurate set of form-form correlations. Rather, following MacWhinney (1975, 1978, 1982, 1984) and Braine (1976, this volume), the child appears to be guided by two principles in deciding what to correlate with what. One principle is that of semantic connectedness. The other is positional patterning. Together, these two principles tightly delimit the scope of the co-occurence patterns that the learner considers. At the same time, by examining formal correlations between items that are positionally connected and semantically related, the learner can acquire the basic form-form correlations of the language. The fact that children have no trouble picking up discontinuous forms such as "call . . . up" as well as long-distance dependencies such as subject-verb agreement should not be interpreted as indicating that all correlations between all forms are tracked all the time. For such discontinuous forms there are also continuous versions in which the two parts of the construction occur next to each other (e.g. "call up" or "we are") and these nondiscontinuous variants can serve as initial guides to the formation of the systems of horizontal and vertical correlations.

It is important to remember that, although the system is capable of acquiring a complex set of horizontal correlations, the mappings that drive the system are the vertical correlations. Horizontal correlations are acquired in the service of supporting the system of vertical correlations. Vertical correlations are mappings between devices and the functions that they cue and the acquisition of these mappings is driven by the principle of cue validity (MacWhinney et al., 1985). By relating our principle of cue validity to this view of horizontal and vertical correlations, we have moved toward an integration of the functionalism of Bates and MacWhinney (1982) with the correlational learning of Maratsos (1982).

Coalitions as prototypes. In natural languages, mappings of a single form onto a single function are quite rare. Rather, languages make extensive use of polysemy, thereby producing grammatical systems in which the same form can map onto several functions, while the same function can map onto several forms. Taken together, these many-to-many mappings comprise a series of subsystems which we refer to as *coalitions*. The paradigm case of a coalition, exploited in most of our experimental work, is the organization of "sentence subject." In our view, "subject" is neither a single symbol nor a unitary category. Rather, it is a coalition of many-to-many mappings between the level of form (e.g., nominative case marking, preverbal position, agreement with the verb in person and number) and the level of function (e.g., agent of a transitive action, topic of an ongoing discourse, perspective of the speaker). Notice that the entries at the level of form include both "obligatory" and "defining" devices such as subject-verb agreement and "optional" correlates such as the tendency for subjects to be marked with definite articles. This is precisely what we mean when we argue that there is no sharp line between obligatory rules and probabilistic tendencies.

When we say that a mature speaker "knows" the set of connections of the subject coalition, we mean that he "knows" the internal composition of a **prototypic subject** in his native language. It is the system of horizontally and vertically weighted correlations that underlies our view of grammatical categories as prototypes (Bates and MacWhinney, 1982). All of the critical predictions of prototype theory follow from this claim.

- *Family resemblance.* In less-than-ideal communicative situations (where the proposition to be expressed does not contain a prototypic subject), the set of surface forms that comprise "subject" in this language will be assigned by family resemblance, i.e., by "best fit" or "maximum overlap" with the prototypic subject. Hence membership in the subject category is a matter of degree.
- *Heterogeneous membership.* A grammatical category will contain members that overlap with the prototype but not with one another (e.g., a subject that is a non-topicalized agent versus a subject which is a non-agentive topic).
- *Maximum distance from other categories.* Grammatical knowledge involves reciprocal relations among neighborhoods of categories, where category assignment is the joint product of *maximum overlap* with the category that is ultimately assigned and *minimum overlap* with competing categories that could have been assigned. Hence assignment of the subject role involves not only a calculation of goodness-of-fit to a prototypic subject but also a calculation of goodness-of-fit to other grammatical categories (e.g., prototypic object).

Through generations of experience with the competition process, languages have evolved to exploit natural coalitions whenever it is possible to do so. Nevertheless, it does occasionally happen that functions which prototypically "go together" have to be split apart and assigned to different items in order to express an idea adequately. To illustrate, consider what happens when the coalition between agency and topicality breaks down in English and Italian. This can occur, for example, when we need to topicalize "the ball" even though "John" did the hitting. In such cases the grammar has to determine which of the two elements should "win" access to devices like preverbal positioning and verb agreement. We have classified solutions to this problem into two basic types: "compromise" or "divide the spoils."

A typical compromise solution, provided by both English and Italian, is selection of the passive; the patient/topic "ball" wins access to the major subject devices, but the agent is placed in a special "by clause" that signals its continued semantic role. In terms of prototype theory, this is the kind of "hedging" and category-mixing that often occurs when categorization decisions have to be made for peripheral members. In a sense, it is a sentence-level analogue to word-level expressions like "an ostrich really is a bird," designed to mark explicitly the peripheral status of a category assignment.

Topicalization is an illustration of a typical "divide the spoils" solution. In topicalization in Italian, preverbal position is assigned to the topicalized patient, but verb agreement is still assigned to the agent. In other words, the set of surface devices comprising "subject" is simply split and assigned to separate elements. This kind of splitting rarely occurs in English, although we do find informal constructions like "Now that I'd really love to see!" The highly-correlated subject devices in English tend to be assigned as a block, while the lower correlations among the same devices in Italian permit the coalition to be split up for non-prototypical situations.

The point is that a series of compromises are made in both sentence comprehension and sentence production. The ideal situation does not always hold. In fact, the fully prototypical instance of a category such as "subject" may actually be fairly rare (like the "ideal member" that is extracted but never taught in studies of artificial category learning such as Posner and Keele, 1968). This is possible because our knowledge of a "prototypic subject" is the emergent property of a great many weightings between individual forms and functions. It is the result of a lifetime of distributional analysis and not a template derived from any single instance of grammatical learning.

Competition. The model assumes dynamic control of the mapping of form onto function in comprehension, and the mapping of function onto form in production. This mapping is understood to be governed by a system of parallel activation with strength-based conflict resolution much like that found in word-level processing models such as Thibadeau, Just and Carpenter (1982) or McClelland and Rumelhart (1981). The competition model extends these word-based models to the sentential level to account for assignment to grammatical roles and other parsing decisions in comprehension.

To illustrate what we mean, consider the kind of word-recognition system modelled by McClelland and Rumelhart to account for experimental findings by Glushko (1979). People know how to pronounce a "new" or non-existent string like *mave,* even though they have never seen it before. Usually they will pronounce it to rhyme with *cave,* but occasionally they will pronounce it to rhyme with *have.* Glushko suggested that speakers make their decisions *not* by applying an abstract set of phonological rules but by a process of analogy. Specifically, when the letter string *mave* appears, all of the existing words that overlap partially with this nonsense string are activated simultaneously. Each of the real world candidates has a basic activation level reflecting (at least in part) its baseline frequency in the language. The decision on how to pronounce this nonsense input and the time taken to reach that decision emerge out of the competition among all of these partially overlapping "demons": *have* as well as *cave, save, rave, mane, mate, cane,* etc. In the case of *mave,* the high frequency candidate *have* does occasionally win out, but it is usually overwhelmed by the greater number of word candidates with a long "a" pronunciation. Hence deci-

sions are a combined product of the number of different types in the competition pool, and the activation weights associated with each type.

A syntactic analogue to this process can account for a number of robust phenomena in our cross-linguistic sentence comprehension data. In a series of studies with children and adults, we have presented listeners with simple sentences comprised of two concrete nouns and a transitive action verb. The sentences always represent some orthogonal combination of lexical semantic, pragmatic, grammatical and/or phonological cues: different word orders (NVN, VNN, NNV), presence of contrastive stress or presence of topicalization information for one of the two nouns, subject-verb agreement, case marking, etc. In some of our studies in some languages, the resulting list of stimuli includes a mixture of grammatical and "semi-grammatical" sentences, e.g.,

- The dog are kicking the cows.
- The pencil the horse is pushing.
- Is eating the ELEPHANT the tiger.

As we will present in more detail below, there are massive differences between languages in the way that listeners respond to such stimuli. Most of these differences follow directly from calculations of cue validity as described above. For example, Italians make greater use of both semantic contrasts and subject-verb agreement than their English counterparts; English listeners make greater use of word order than any other cue.

In addition, however, we have also uncovered some interesting new information about the specific strategies used in different languages to deal with combinations of word order and stress information. For example, English listeners have not only an overwhelming SVO strategy to deal with NVN sentences, but they also have very strong and reliable VOS and OSV strategies to deal with the two respective non-canonical word order types. Italians have a rather weak SVO bias for NVN stimuli. But they have no bias at all for NNV or VNN word orders. In fact, Italians can make use of word order information only in the presence of certain characteristic order/stress configurations: SVO applies consistently only if the NVN sentence has default stress; a weak but reliable SOV bias appears if the second noun is stressed in an NNV; a weak but reliable VOS bias is applied only if the first noun is stressed. For every language that we have looked at, we have discovered some set of word order biases that do not follow in any straightforward or obvious way from known facts about basic word order types in the language. That is, the word order biases cannot be attributed to any single word order "model."

However, we discovered that word order and/or order/stress biases *can* be accounted for by the parallel activation of *all* the partially-overlapping word order types in the language. That is, if a syntactic analogue to the *mave* example is going on, we should get exactly the results that we have obtained in all of the

languages examined so far. Suppose, for example, that we line up all the possible syntactic patterns for talking about "John hit the ball" in English. Ignoring morphology and looking only at the "islands" of constituent ordering (e.g., treating a cleft sentence like "The one who *hit* the *ball* was *John*" as a VOS), it is clear that SVO is the statistically predominant ordering in English, followed by OSV and VOS. Carrying out a similar exercise in Italian, we find a much weaker bias toward SVO with essentially random probabilities for SOV versus OSV, VOS versus VSO—unless we take stress into account. For French, alternative word orders are possible only in the presence of a complex set of clitic markers. And yet, in our experiments, French listeners behave almost exactly like Italians, distrusting word order and making decisions primarily on the basis of semantic and morphological cues even though, in our sentence stimuli, we did not provide clitics to "release" word order variation. Their behavior with respect to word order alternatives makes sense only if they are engaged in some kind of competition process that involves a competition among *partially* as well as *completely* overlapping word order candidates. That is, the French listener's behavior is influenced by partially activated cliticized phrase structures, even when no clitics appear in the input.

Ongoing updating. In order to control the real-time interaction of the various cues participating in the competition, we believe that the parsing system engages in an ongoing updating of assignments of nouns to case roles. For example, when parsing a sentence such as "The dogs are chasing the cat," the assignment of "dogs" as the agent is first promoted by its appearance as the initial noun. Then the fact that "are chasing" agrees with "dogs" in number further supports this assignment. Finally, when the singular noun "cat" appears post-verbally, its binding to the object case role further supports the candidacy of "dogs" as the agent. Thus, at each point in sentence processing the mapping from the lexical item "dogs" to the agent role is updated. In this particular case, each updating increased the strength of this assignment. In other cases—particularly in languages that permit a great deal of word order variation—assignments may wax and wane in strength across the course of sentence processing.

Having reviewed the basic claims of the model, let us note some ways in which the model relates to some older models of the learning process. Within our model, language learning is viewed as a process of acquiring coalitions of form-function mappings, and adjusting the weight of each mapping until it provides an optimal fit to the processing environment. This is quite similar to the process that Gibson (1966) describes as "detection of invariance" and/or "acquired distinctiveness of cues." Remember, however, that it is the organism who determines the set of functions to be mapped onto forms. In this sense, then, our connectionist model has little in common with passive associationist theories of the Hullian variety. Instead, it might be more appropriate to label this kind of learning theory as a "neo-Tolmanian" approach, insofar as (a) the goals and

expectations of the organism play a major role in determining what will be learned, and (b) form-form, function-function, and form-function connections can be observed and pondered in their own right whether or not the organism is currently being driven to meet some primary need.

Evidence for Cue Validity in Acquisition

The major developmental prediction of the Competition Model is that cue validity will determine the order in which grammatical devices are acquired. Cues differ widely in their validity across languages, and the model predicts that these differences in cue validity will be reflected in differences in the course of use of the cues in language acquisition. The prediction is that children should be sensitive from the beginning to the information value of particular perceptual patterns, and will go to work first in those forms that promise a greater "payoff." This is most assuredly a **minimalist claim,** i.e., a hypothesis that is offered with the hope and conviction that it will be falsified. It is, however, a claim that accounts for a great deal of variation in language learning across natural languages.

To illustrate, let us consider two would-be universals that have been disproven by cross-linguistic evidence:

1. Semantic cues to agent-object relations will be acquired before word order cues.
2. Word order cues to agent-object relations (and to other sentence roles) will be acquired before grammatical morphology.

Semantics before word order.

Bever (1970) proposed that three-year-old children will rely primarily on semantic strategies in comprehending simple declarative sentences, whereas four year-olds make primarily use of word order. A related "semantics first" hypothesis has been offered by Strohner and Nelson (1974). Chapman and Kohn (1978) provided some qualifications on this proposal differentiating between abstract semantic contrasts like "agent/patient" and "probable event" relationships between the protagonists in certain well-known events (e.g., mother and baby in "The mother feeds the baby"). They suggested that probable event strategies would precede the application of word order principles, but abstract semantics might be rather late. In either case, however, the main idea is that some kind of semantic information would be used by all children in all languages in advance of the grammatical information in word order.

Although the "semantics first" hypothesis has gained rather wide acceptance, it has not held up under cross-linguistic tests. In French, Sinclair and Bronckart (1972) found that reliance on SVO word order increased from two- to seven-years of age. But interpretations based on event probabilities were also

strong at all ages. In English, SVO word order eventually dominates over any alternative strategy. In contrast, at no point in this study of French children did SVO word order "defeat" a probable event strategy when the two were placed in competition. A recent study by Hakuta (1982) provides still clearer evidence regarding the status of lexical semantic strategies in languages other than English. Hakuta's stimuli were created by random assignment of nouns such as "tiger," "goat," "chair" and "banana" to verbs like "push" and "chase." When word order and lexical semantics were placed in direct competition in Japanese, lexical semantics dominated quite clearly.

Of course, when data are presented for only one language, we can always argue that the results are stimulus-specific, and probably would have created the same effects in *any* language. For this reason, we have tested the relationship between abstract lexical semantics and word order in several different languages, looking at adults as well as children to determine the "end state" of development in each language.

First, notice a certain anglocentric bias in the previous experiments on children, i.e., the assumption that word order *ought to* dominate in a mature listener. In the competition model, there is no pressure toward the licensing of such biases. Instead, the relative strengths of word order and lexical cues are a function of the relative validity or information value of those cues in each language. In languages like Italian (as well as French and Japanese), word order variation for pragmatic purposes is very common. A mature listener would do well to "trust" semantics over word order in many situations. And that is exactly what we find in studies of adults. In all of these languages, an abstract animacy contrast between two nouns wins out over word order relations when the two pieces of information are placed in direct competition. This includes studies of Italian (Bates, McNew, MacWhinney, Devescovi & Smith, 1982; MacWhinney, Bates & Kliegl, 1984), French (Kail & Combier, 1983), and Japanese (Ito, in press), as well as German (MacWhinney, Bates & Kliegl, 1984), Serbo-Croatian (Smith & Mimica, 1984), Mandarin Chinese (Xiao-Chun, 1981), and several other languages—although the magnitude of this semantic bias varies from one language to another in accord with language-specific differences in the cue validity of word order information. It actually seems fair to conclude at this point that English is an exotic language with a word order bias that has so far failed to appear with equivalent strength in any other language.

If cue validity is a major determinant of the acquisition of forms across languages and if the same or equivalent stimuli are used in each language, then we should expect to find an early SVO word order bias in English children, contrasting with an early semantic bias in languages like Italian or Serbo-Croatian. We recently tested this hypothesis with English and Italian children between two and seven years of age (Bates et al., 1984). The stimuli, like those used in studies of adults, were all simple, active, declarative sentences with two concrete nouns and a transitive action verb. All three word orders were tested (NVN,

NNV, VNN) in orthogonal combinations of animacy (first noun animate and second inanimate; first noun inanimate and second animate; both nouns animate) and contrastive stress (first noun stressed; second noun stressed; default stressed only). The results were clearcut: SVO word order was the first cue to have a significant effect on sentence interpretation for English children (starting at two years of age), and animacy was the first cue to have a significant impact on Italians (again, at two years of age). At no point in the English data did an animacy strategy dominate over word order; at no point in the Italian data did word order dominate over animacy. Instead, children tended to veer closer and closer to the adult end point with more and more variance accounted for by the most valid cue in the child's language. There were certain "minor" deviations from this "major" trend, to be discussed in more detail below. In general, however, the data provide a remarkably clear confirmation of the role of cue validity in language acquisition. Very similar findings, with appropriate language-specific variations in timing and magnitude, have also been reported for French (Kail & Combier, 1983), Hungarian (MacWhinney, Pleh & Bates, in press), Mandarin Chinese (Miao, 1981) and Serbo-Croatian (Smith, personal communication).

Word order before morphology. Pinker (1982) offered this proposed universal in this form:

> For case-inflected languages, children will utter sentences in the dominant word order, and will use the dominant word order as a cue in comprehending sentences, before they have mastered their language's morphology.

Around the same time, however, Slobin and Bever (1982) offered a rather clear counter-example to Pinker's claim, in data on sentence interpretation by Turkish children. The Turkish system of case inflections is perhaps the most regular, semantically transparent, and unambiguous system in the world. In our terminology, the cue validity of nominative/accusative case marking as a guide to semantic relations approaches 100%, and hence should be the first cue adopted by Turkish children. Indeed, Slobin and Bever showed that Turkish children have completely mastered the use of these case contrasts by two years of age; by contrast, they show little or no sensitivity to word order contrasts until around the age of four, and even then they apply word order only to "aberrant" sentences in which case information is ambiguous. Independent information on language production suggests that these contrasts also appear very early in expressive language, perhaps even in the one word stage (thus providing a considerable complication for our notion of a "single word utterance").

Gleitman and Wanner (1982) have provided a counter-argument to Slobin and Bever, in an effort to salvage the "word order first" hypothesis. They note that the case markers of Turkish can receive full stress, very much like a free-

standing lexical item. They do not, then, qualify as "clitics," i.e. as word- and phrase-building morphemes that can neither stand alone nor receive full stress. If children are innately endowed to perceive only unstressed clitic forms as elements of grammatical morphology, then they would end up treating Turkish case inflections as full lexical items. As such, the apparent precocity of case morphology in Turkish would not really constitute a counter-example to the "word order first" hypothesis.

However, more recent information on the acquisition of Hungarian (Pleh, 1981; MacWhinney et al. 1985) and Polish (Weist & Konieczna, 1985) suggests that children can also acquire unstressed case inflections before they demonstrate systematic use of word order patterns, in comprehension or production. To be sure, these children do not seem to acquire case with the lightning speed evidenced by their Turkish counterparts—nor would we expect them to, since the Polish and Hungarian case systems are somewhat more irregular than Turkish. However, they use morphological marking before they use word order, unlike children acquiring the nightmarish inflectional systems of Russian (Gvozdev, 1961) or Serbo-Croatian (Radulovic, 1975).

To summarize, the relative timing of semantic, syntactic and morphological developments is at least a partial function of cue validity. There is no universal schedule that determines the order in which children will acquire these surface forms. However, we cannot conclude that order of acquisition is determined *entirely* by the information value of cues. There are a number of other processing constraints that interfere. This brings us to the important issue of **constraints on cue validity.**

Before we proceed, we need to make a few statements about our philosophy of science. We do not consider the competition model a theory. A theory is a set of inter-related hypotheses that can be directly tested and rejected by some line of evidence. They can be found in areas such as particle physics and cellular genetics. In areas such as psychology, we must generally be satisfied not with theories, but with models such as the competition model. A model has much less internal coherence, insofar as it reflects an open-ended or "bottom up" attempt to describe or simulate aspects of the world. For this reason, it cannot be falsified in the strict sense; it can only be confirmed or disconfirmed in pieces. We can always "fix" a model by adding new constraints and principles to account for new findings and by dropping specific claims that cannot stand up to the evidence. How, then, do we ever decide that a model should be rejected? We suggest that models fail to be useful when they become circular, that is when new principles or constraints are added just to save the old ones, with no independent justification. When a model undergoes too many *ex post facto* repair attempts, it finally becomes a patchwork of assumptions that has no architectural center. When this architectural inadequacy becomes evident to all, the model collapses.

In the next section, we would like to point out some places where the minimalist principle of cue validity has failed to account for cross-linguistic variation.

In each case, these failures have led us to add new constraints to the model. Each of the new constraints that we have added so far does have a "life of its own", leading to a new and different set of empirical tests. We will discuss two classes of constraints here: **functional readiness** and **cue cost.**

The Importance of Functional Readiness

Cue validity apparently does play a major role in determining the sequence of acquisition. However, in our studies to date we have found some interesting and consistent exceptions. These exceptions arise when the language makes use of a form to express a function that the children have not yet acquired. In such cases we can say that the function itself is not "ready." This occurs most obviously in those areas where functions play roles that are determined by complex discourse structures. If the child has not yet developed sufficient control of narrative and expository styles to understand the uses of these functions, it will be difficult for her to learn to make correct use of the forms that express these functions.

One set of patterns which is particularly delayed in acquisition are the kinds of secondary word order stress patterns we discussed earlier. These patterns include

- VOS and OSV "second noun strategies" in English
- word order and stress configurations in Italian

Although these patterns are quite reliable in adults, showing up in several different studies, Italian and English children show no apparent use of these adult patterns until somewhere between seven and nine years of age. In fact, children in both languages show a small but reliable tendency in the first years to choose the stressed element as the subject—exactly the opposite of the adult pattern. We have suggested elsewhere (MacWhinney & Price, 1980) that in this case children are using an essentially non-linguistic strategy: if you don't know what else to do, choose the noun that the experimenter said loudest. This pattern of results indicates that children can certainly hear the stress manipulation. We believe that they fail to make consistent use of them because they are not yet ready to acquire the underlying functions.

Because they cannot make effective use of these patterns, Italian children develop a secondary strategy of their own. They generalize the first noun strategy derived from SVO to other word order types. This strategy starts between four and five years of age and does not drop out until after the age of nine when we finally see reliable use of stress to interpret NNV and VNN sentences. During this period of "word order overgeneralization," Italian children make *more* use of word order and *less* use of animacy than their adult counterparts (although word order never actually "wins out" over lexical semantics). One way of understanding this overgeneralization is to say that children are overgeneralizing

the pattern they do understand rather than making use of patterns whose function they do not understand.

Kail and Combier (1983) and Kail and Hernandez-Charvillat (1983) have obtained almost identical results in a comparison of children and adults in French. A much weaker version of "first noun overgeneralization" has also been noticed in Hungarian and Serbo-Croatian (Pleh, 1981c; MacWhinney et al., 1985; Slobin and Bever, 1982; Smith, personal communication). In these studies the overgeneralization of first noun choice also starts around four years of age, although it drops out considerably earlier than it does in Italian or French. Similar results can be found in Miao's (1981) study of Chinese children. And a very small bias to attend to word order seems to get off the ground around four years of age in Slobin and Bever's (1982) study of Turkish.

No such overgeneralization occurs in our data on English children. Other investigators have reported a tendency for English children to generalize SVO strategies to the passive—a strategy which, interestingly, peaks between ages four and five. But the first noun bias is apparently not extended to NNV and VNN constructions in our research. Bates et al. (1984) suggest that the first noun tendency is somehow *blocked* in English, because the English children are sensitive at some level to the factors that create VOS and OSV biases in adults. We will return shortly to an explanation of how this might occur.

In any case, we do seem to be confronted here with a consistent cross-linguistic exception to the role of cue validity in determining acquisition of word order and other grammatical forms. To the extent that these patterns of "first noun dominance" vary markedly in size, range, and time of offset, we would not want to argue that they are universal. And it is possible that they reflect very different principles from one language to another. For example, Bever (1975) has reported that first noun strategies in the interpretation of the passive are correlated with degree of left-hemisphere bias in a dichotic listening task, among a sample of English four- to six-year-olds. Pleh (1981) carried out a similar study in Hungarian and found diametrically opposed results: to the extent that a child imposed a word order strategy in interpreting case-marked sentences in Hungarian, that child was significantly *less* likely to show a left-hemisphere bias on a dichotic listening task. Both authors suggest that degree of lateralization can be taken as an index of linguistic maturity. If they are correct, then word order biases are *more* mature in English and *less* mature in Hungarian. Nevertheless, the fact that they can exist at all needs to be explained in a performance model like the Competition Model.

To explain these phenomena we invoke a principle called **functional readiness.** Functional readiness means that children will not acquire a complex form until they can assimilate it, directly or indirectly, to an underlying function. This principle was used by Brown (1973), MacWhinney (1975), and Slobin (1973) to account for the acquisition of grammatical markers. However, here we are extending it to account for aspects of syntactic development. Following MacWhin-

ney (1978) we can distinguish between the acquisition of forms, the acquisition of functions, and form-function mappings.

- **The acquisition of functions:** When the child uses and perceives new concepts and patterns, these new patterns become integrated into mental processing as new functions. For example, after a visit to a farm, a child may begin to develop the concept of a "cow." This concept is then a function which the child may seek to express.
- **The acquisition of forms:** A form is assimilated without being bound to a function when the child does not understand what it is for, but learns by rote that it is associated with other forms that he does understand and wants to use (e.g., the morphology associated with gender may be picked up by necessity as the child acquires lexical items that are marked for gender).
- **Form-function mappings:** By relating forms to functions, the child ties a function to a form, and thus gives motivation for that form. For example, when the child learns that subject-verb agreement signals topic/agent roles, she has acquired a new form-function mapping.

Karmiloff-Smith (1982) has argued that children do not acquire the inter-sentential functions of surface forms until seven years of age or older. Her conclusions are based upon a variety of facts about the way that children use and talk about pronouns and determiners in French and English. If Karmiloff-Smith is correct, then the same principle could account for delays in the acquisition of secondary word order patterns and order/stress configurations in our data. The analysis we are proposing can be outlined as follows:

- In adults, the application of these secondary strategies depends upon the presence of complex and low frequency phrase structure types in the "competition pool."
- These phrase structure types are governed by underlying discourse cohesion functions that span several propositions.
- Young children do not yet understand these inter-sentential discourse functions.
- Therefore, they cannot assimilate the complex phrase structure configurations directly (i.e., they cannot map them on to appropriate functions).
- They also fail to pick up complex syntactic patterns by indirect assimilation, i.e., by rote association with things that they do understand, perhaps because they are not exposed to enough exemplars.
- As a result, the process of sentence comprehension in very young children involves competition among a much smaller pool of sentence patterns.
- The small competition pool can result in temporary patterns of first-noun overgeneralization. These are similar to the temporary overgeneralizations

observed in the acquisition of morphology (see Rumelhart and McClelland, this volume). They peak when a certain proportion of subject-initial phrase structures have accumulated and drop again as further evidence comes in.

The last point leads to the prediction that similar overgeneralizations will be observed in comprehension and production. We do not yet know whether this is the case. It is certainly true that preschool children use only a subset of the possible syntactic patterns in their language. However, we need more evidence about such things as the range and frequency of order/stress configurations in child speakers in "free word order" languages like Italian and Hungarian (cf. Bates, 1976).

We also need to explain why English children fail to overgeneralize SVO to the same extent as children in the other languages studied to date. If they do not have the phrase structures responsible for the OSV and VOS biases of English adults, then why should any "blocking" occur at all? We may have to postulate a "two-tiered" membership in the competition pool of either word or phrase structure candidates. Active members in the pool are members that have been functionally assimilated (as defined above). These candidates exert the greatest force in a parsing decision. However, the child may well retain some memory of sentence types that he heard but failed to understand. If enough of these exceptions accumulate, they may serve to block certain generalizations that would be possible in their absence. This "second tier" notion is similar to the "waiting room" idea discussed by Ammom and Slobin (1979) and to the "file of unknown forms" in MacWhinney (1978).

The functional readiness principle can handle a variety of exceptions to cue validity. It is really a very simple notion, tantamount to "What you don't know can't hurt you." Furthermore, applied to the contrast between sentence-level and discourse-level functions, the principle has a certain amount of independent justification (e.g., the research by Karmiloff-Smith). But just how powerful is this principle? Are there findings that it *cannot* explain? In a word, yes. In the next section we turn to some exceptions that require a very different kind of explanation (if they are to be explained at all).

The Role of Cue Cost

The concept of cue cost did not originate with us. It is discussed by Carroll (1978) as a source of constraints on cue validity in sentence interpretation. Simply stated, a very informative cue may be used less than we would otherwise expect, because the processing costs involved in using that cue are unusually high. We have found it useful to break this general notion down into two distinct components: perceivability and assignability.

Perceivability. We use the term "perceivability" to refer to the extent to which a listener encounters difficulty in trying to detect a cue for use in sentence processing. Consider, for example, the extreme example of subject-verb agreement in spoken French, e.g., the contrast between

Elle mange	She eats

and

Elles mangent	They-feminine eat

For most verbs in conjugations, the clear-cut written difference between the singular and plural form of the verb is entirely inaudible. Even though the agreement contrast is distributed quite faithfully through written texts of French, it is an imperceptible cue in the oral language. Cue validity means very little if a cue cannot be heard at all.

MacWhinney, Pleh, and Bates (1985) have shown that less drastic differences in the perceivability of cues can also have a significant impact on the way those cues are used in sentence interpretation. In Hungarian, case contrasts are very high in cue validity. However, some case suffixes are easier to hear than others. The nominative/accusative contrast can involve a strong vowel contrast plus the addition of a final /t/ (kutya—kutya't), or it may simply involve the addition of the final /t/ (mokus—mokust). When the /t/ follows a dental or alveolar consonant, it is difficult to identify with certainty. In sentence interpretation experiments with adults and children, this difference in the perceivibility of cues interacted with cue validity in determining the probability that a Hungarian listener will rely on case information. If case competed with other cues (e.g., word order and semantics), the strong vowel form of the suffix quite clearly "won" the competition; but if the weaker consonant form was involved, listeners would often (though not always) take the "conspiracy" of order and semantics into consideration in making their decision. MacWhinney et al. suggest that this is a morphological version of the "phoneme restoration effect" of Warren and Warren (1976), e.g., the tendency to hear a stimulus like "(cough) eel" as "wheel" or "meal", depending on the sentence context. Another way of putting it is that a lifetime of not being sure whether a case contrast is there or not has led listeners to "distrust" that cue, even in those instances when it is perceived.

We are currently pursuing some extensions of the perceivability notion in studies of the acquisition of clitic particles in Italian and French. There are a number of places where a preverbal object marker is equally informative in both these languages. However, the respective French and Italian markers differ in their perceptual salience, depending on such factors as the kind of verb that follows. For example, if the clitic precedes a participial construction, as in the French "Les ont fait" versus the Italian "Li hanno fatto," the fact that a

consonant is pronounced between the clitic and the auxiliary should make the French clitic relatively more salient. We are interested in determining the extent to which subtle phonological factors like these influence the acquisition of clitic forms as a cue to underlying semantic relations. Obviously this complicates any predictions based on cue validity, since a whole catalogue of phonological factors would have to be taken into consideration. But if that is the way things are, so be it. Interactions between phonology and syntax may be hard to handle theoretically, but they are certainly testable.

Assignability. The second class of cue cost factors have to do not with perception but with memory. "Assignability" refers to the ease with which a given cue can be assigned to a role. A cue which can be utilized as soon as it is perceived is maximally high in assignability. In a language like Turkish, with a completely unambiguous case system, case suffixes have this status. The semantic role of a noun is assigned as soon as the suffix on that noun can be classified. This is what Ammon and Slobin (1979) and Johnston and Slobin (1979) refer to as a "local cue." By contrast, a cue which spans two or more disparate and perhaps discontinuous elements can be referred to as a "global cue." These cues are low in assignability, because the listener has to wait until several pieces of information are in before an assignment can be made. If the processing system is under stress and/or if the processor has limited auditory storage, global cues may become so costly to handle that they are abandoned despite their information value.

We have used the principle of assignability to explain another kind of anomaly in our developmental cross-linguistic data. As we have already pointed out, Italian adults use the information in subject-verb agreement in the same way that speakers of case-marked languages use case inflections: as a very powerful, almost deterministic cue to sentence meaning. Agreement is not maximally high in validity, because it is occasionally ambiguous (e.g., when both the subject and the object noun phrase are third person singular or third person plural). However, if cue strength does depend on cue validity to the extent that we believe it does, then reliance on agreement in Italian should be as strong as reliance on case in languages like German and Serbo-Croatian (where case is also ambiguous, to approximately the same extent). The adult data support this interpretation. But the developmental comparison looks very different. Serbo-Croatian children begin to make systematic use of case information by three years of age (Smith, personal communication). And according to some analyses we were allowed to carry out on Slobin and Bever's data, 2 year-old Yugoslavians may already have at least some sensitivity to case marking.

On grounds of cue validity alone, we should expect similar behavior by Italian children in the use of subject-verb agreement. Devescovi and Caselli (1985) tested this hypothesis in study of Italian children between three and nine years of age, varying three factors: word order, animacy, and number agreement between

the verb and one of the two nouns. Results for word order and animacy replicated our findings in previous studies on Italian in every respect. So there is no reason to think that, just this once, we did something wrong. Nevertheless, findings for verb agreement were very surprising. The agreement factor did not become the dominant cue until the age of seven—three to four years after a similar level was reached by Yugoslavian children using the "best" morphological cue in their language. There are no "functional readiness" principles to account for this discrepancy since agreement maps onto basic sentence-level event structures in exactly the same way that case signals meaning in Serbo-Croatian and word order signals meaning in English. Nor can we turn to principles involving the perceivability of cues. The singular/plural contrast on verbs is very salient, involving an additional consonant/vowel syllable on the plural form (e.g., mangia/mangiano). Furthermore, we know on the basis of other data that Italian children mark some forms of verb agreement in their productive language by the age of two, with the singular/plural contrast in the third person coming in at least by the age of three (Bates, 1976). Thus children have this surface form in production; they simply are not using it in comprehension.

These data can be explained if we assume that a global agreement cue places too much load on the young child's limited memory system. To use agreement correctly in the third person, the child must store *1.* the noun that will ultimately agree with the verb, *2.* the verb itself, and *3.* one or more competing noun phrases which could also agree with the verb (thereby rendering agreement information ambiguous and useless). But do we have any independent evidence that the "low assignability" of agreement creates a delay in acquisition? This idea was tested again in a recent study of sentence interpretation in French adults and children (Kail and Combier, 1983; Kail and Hernandez-Charvillat, 1983). These authors used two agreement markers as variables (subject-verb agreement and clitic-object agreement) with word order and animacy. In adults, both agreement markers played a major role in sentence interpretation. But in children, neither of the agreement cues were used to any significant degree until after five years of age. This is, at the very least, a replication and extension of the anomaly posed by Devescovi and Caselli. But it also suggests that the late acquisition of agreement markers is a fairly general phenomenon, in contrast with various studies on the acquisition of case. So the "assignability" hypothesis is rendered more credible.

A final bit of support for the role of assignability comes from a comparison between children and adult aphasics in the same kinds of experiments. We would expect the two cue cost variables (perceivability and assignability) to play a role in language processing in aphasia, since the performance of brain-damaged patients is presumably the reflection of a system under stress. However, there is less reason to expect the developmental principle of functional readiness to affect the performance of brain-damaged adults. Processing may be under stress, but knowledge of how language works—once it is acquired—should not easily go

away. Some of our preliminary cross-linguistic findings in aphasia provide support for this view. For example, the second noun strategies shown by normal adults in English are still detectable in English aphasics, even though such strategies are not apparent at all in the performance of English children under five. However, a comparison of Italian and Serbo-Croatian aphasics shows that the assignability contrast does play a role when the adult system is under stress. Both case and agreement are impaired in these patients relative to normals in the same two languages. However, the "damage" done to case is measurably smaller than the "damage" done to agreement. In other words, "local" morphological cues are more resistant to brain damage than distributed, "global" morphology that is low in assignability.

In short, we think that the two principles of functional readiness and cue cost serve not only to constrain the competition model, but also to enrich it. The two principles are fully in accord with our emphasis on form-function mappings and the role of competition in processing. These principles are non-circular, motivated by independent lines of evidence, and they are testable. The competition model continues to assist us in studying variation in the acquisition process across natural languages by both accounting for old data and stimulating new research. However, we are less sure how the model will stand up to data on individual differences within a single language, which we will now confront.

INDIVIDUAL DIFFERENCES

The literature on individual styles of language acquisition has grown considerably in the last decade, although the results pertain almost exclusively to English. Our purpose here is simply to indicate the *kinds* of variation that have to be accounted for by acquisition theories. The reader is referred elsewhere for more detailed reviews (Kempler, 1980; Nelson, 1981; Bates, Bretherton, & Snyder, in press; Bretherton et al., 1983).

At the lexical level, the parameters of variation were first set out by Nelson (1973), in a study of the first 50 words acquired by 18 children in the second year of life. She described a dimension of variation ranging from *referential style* (i.e., vocabularies with a high proportion of common object names) to *expressive style* (i.e., heterogeneous vocabularies containing items from a variety of form classes including some frozen phrases like "Stop it" and "I love you."). The term "referential" is based on the tendency for children at this end of the continuum to focus on the process of object naming. The term "expressive" captures the tendency for children at the other extreme to focus more on the social/regulatory uses of language. This dimension was quantified through simple proportion scores (total object names divided by total vocabulary), resulting in a normal distribution with most children falling somewhere in the middle. In

other words, this is not a bimodal categorization of children into distinct language types. Nevertheless, for children at the extreme ends of the continuum, linguistic behavior looks so strikingly different that it seems to reflect the operation of qualitatively different acquisition mechanisms and/or linguistic environments. The range of variation observed in Nelson's study has now been replicated many times over. Furthermore, Snyder, Bates and Bretherton (1981) have shown that the referential/expressive split occurs in comprehension as well as production, beginning as early as 13 months of age when vocabularies average only 10–12 words. This is, then, a very general phenomenon that spans the period from first words to sentences.

At the grammatical level, the first major report on individual differences was offered by Bloom, Lightbown and Hood (1975). In a longitudinal study of four children, these authors introduced a contrast between *nominal style* (i.e., multiword constructions composed primarily of nouns and other content words), and *pronominal style* (i.e., multiword constructions in which the same meanings are conveyed with non-specific pronominal forms). All of the children in the Bloom et al. study produced some "pivot-open" constructions like NO + X, MORE + X, or ANOTHER + X. However, to encode semantic relations involving action, location and possession, two of the children produced "telegraphic" combinations of content words: KATHRYN SOCK, TOUCH MILK, etc. In particular, they tended to refer to themselves by name. To express the same meanings, the other two children were more likely to use non-specific forms of reference: I finish, my truck play it. In particular, these children were more likely to refer to themselves with first person pronouns. The terms "nominal" and "pronominal" are based on the striking difference between the two types in first person reference. Note, however, that this particular difference occurred only in the first stages of grammatical development; by the time all four children had attained a Mean Length of Utterance of 2.5 morphemes, contrasts in noun and pronoun use had disappeared.

The nominal/pronominal dimension has been replicated in several other studies of English children, including a large-sample study by Bretherton et al. (1983). These investigators also showed that the noun/pronoun distinction is only one manifestation of a more general contrast involving presence or absence of grammatical morphology during Stage I. Children at the extreme nominal end of the continuum produce the kinds of uninflected telegrams that were once believed to constitute a universal starting point for grammatical development. However, children at the other end never pass through a telegraphic stage; their first word combinations are already heavily inflected with a large number of free standing function words. Like referential/expressive style, this nominal/morphological dimension is normally distributed. That is, most children cluster around a midpoint in the distribution, using a mixture of sentence types. For example, it is quite common for children in Stages I and II to vacillate between two or more forms of first person reference (e.g., Bates' 23 month old daughter, who pro-

duced the single utterance "I do that. . . . Julia do that . . . me do that"). Again, however, children at the extreme ends of the distribution look so different from one another that it is hard to believe that they have the same "theory" of their native language.

Although the nominal/morphological contrast itself has a short half-life, we now know that it is linked to other dimensions of variation throughout language acquisition. For example, there is a connection between referential/expressive style in the one word stage, and nominal/morphological contrasts in early multiword speech. As one might expect, referential children are more likely to start their grammatical careers with telegraphic speech; expressive children, who have heterogeneous vocabularies from the very beginning, also have heterogeneous mixes of content and function words in their very first sentences. Furthermore, Horgan (1981) suggests that the nominal/referential children are "noun lovers" who continue to emphasize the development of noun phrases and their associated morphology in the more advanced stages of language acquisition; by contrast, pronominal/expressive children can be characterized as "noun leavers," who spend more time elaborating verb phrase morphology across the following months and years of language learning.

There are, then, stylistic links between lexical and morphological development throughout the early stages of language acquisition. Other evidence suggests that the same stylistic dimensions can be found in the development of phonology. It has been noted for some time that pronominal/expressive style is associated with a greater tendency to imitate at both the word and sentence levels. But the imitation tends to be imprecise, focussing on the suprasegmental/intonational "packaging" of a target rather than its phonetic details. In fact, several studies of lexical and grammatical development have noted that the speech of pronominal/expressive children is difficult to understand and transcribe—a tendency that Horgan refers to as "mushmouth." Researchers at the Stanford Child Phonology Project have tried to explicate this notion of intelligibility. Researchers at the Stanford Child Phonology Project (Vihman, 1985; Vihman, Ferguson, & Elbert, in press) have tried to explicate the notions of intelligibility and phonological style. At one year of age, style revolves around the number and identifiability of consonants in babbling and speech. This measure is associated with phonological consistency at age 3, i.e., whether a child pronounces a given word in the same way across instances. Most important, the dimension of *phonological* consistency at 3 is significantly correlated with *syntactic* consistency in the same children. In other words, the same mechanisms may be responsible for individual differences in phonology and grammar.

There is a great irony in these findings on consistency and variation in phonology and grammar. As Goldfield and Snow (in press) point out, the language acquisition literature of the 1960's left us with the impression that all children pass through a universal course of development from naming, to telegraphic speech, to grammar. We know that this is, at best, an optional route to language

acquisition. Why did we miss the other half of the story for so long? In the first classic longitudinal studies of language acquisition, there was absolutely no reason to believe that variation in phonology would bear any relation at all to semantic or grammatical development. Children were chosen for inclusion in the studies on the basis of several criteria, including the relative intelligibility of their speech (Bellugi, personal communication). The purpose of this selection criterion was obviously to facilitate the lengthy process of tape transcription. But it also influenced our theories of language acquisition for at least a decade.

A number of studies are now underway to determine whether linguistic variation is correlated with dimensions outside of language proper: cognitive styles in problem solving and symbolic play, aspects of personality or temperament, and a host of social/environmental factors including social class, birth order, and maternal style. For example, preliminary findings suggest that children at the pronominal/expressive end of the continuum are a bit more sociable, and that they are interested in trying to act and sound like other people as much and as often as possible (a tendency that is frequently manifested in immediate imitation of adult speech). There are also differences in the way that adults respond to these contrasts in style, reactions which can either augment or damp an initial tendency by the child to follow a particular path into language acquisition (Goldfield and Snow, in press). Nevertheless, the differences that have been observed in linguistic input—for some children some of the time—do not seem to be sufficient to account for the robust and reliable patterns of variation that result. For example, if we factor social/environmental variables out of the language data, correlations among variables in the referential/expressive "strands" of language development are relatively unaffected (Bates et al., 1985). We do not know whether these strands of variation are unique to language, or whether they represent linguistic versions of a more general cognitive style. However, we can conclude that these differences originate in the child.

CAN OUR MODELS DEAL WITH THESE DATA?

The data on individual variation mentioned above constitute a major challenge to current models of language acquisition. For one thing, it is difficult to defend a universal and stage-like passage from reference, to predication, to grammar. Certainly, naming does precede *productive* control over grammatical rules, even in children at the extreme pronominal/expressive of the continuum. In fact, Bates et al. (1985) have presented evidence suggesting that rate of lexical development at one year of age is an excellent predictor of grammatical development in the third year. But the passage is apparently neither simple nor linear. For example, we might have expected to find that children who start out early in collecting and using inflections and function words would have a headstart in the analysis and mastery of grammatical morphology. However, Bates et al. report

that the ratio of closed class elements in the child's speech at 20 months is significantly and *negatively* correlated with the same measure at 28 months. In other words, the children who ultimately make more progress in the use of grammatical morphology are the referential/nominal/telegraphic children who studiously leave those elements *out* of their speech during the phase of first word combinations. This suggests, among other things, that the use of grammatical morphology in Stage I is based primarily on rote, unanalyzed expressions and not on "rules" of any sort. In any case, the first appearance of a grammatical form is in no way a harbinger of its eventual mastery. At least at first glance, these nonlinear and contradictory patterns of acquisition seem to present problems for a data-driven, connectionist model of language learning—particularly if, as seems to be the case, these variations are not caused by differences in the language environment. From the point of view of the competition model, successive stages in language acquisition should be shaped by the information value of form/function mappings in the linguistic input—modulated, as we have noted, by such factors as functional readiness and cue cost. As the model is presently constituted, there are no mechanisms for predicting these consistent patterns of individual variation.

The individual differences data also provide little comfort for proponents of a nativist, modular theory of language acquisition. If parameters are set by "input triggers," then the intermediate stages of acquisition should look the same for all children within a given language—assuming that the linguistic environment is in fact the same on all the relevant dimensions.

One possible solution for a modular theory is to argue that distinct components of language processing can mature at different rates. Hence the different profiles that we obtain across children exposed to the same language would derive from an asynchrony between modules in the parameter-setting process. In fact, we suggest that this is the most workable solution to the problem of variation within a language: individual differences are brought about by the differential strength and/or differential timing of two or more mechanisms for processing and learning language. The question is, what kinds of modules can be said to underlie the dissociations observed in early language development? Most of the existing nativist models would predict dissociations corresponding to independent levels of linguistic analysis: phonology, syntax, morphology, and the lexicon. These are the "vertical faculties" discussed by Fodor (1983) in his treatise on modularity, i.e., faculties that have evolved to deal with particular kinds of content or data types. These can be contrasted with "horizontal faculties," i.e., mechanisms like short-term memory or reasoning ability that cut across content domains.

However, as we have already pointed out, the dissociations observed in early language do not respect the boundaries of traditional "vertical modules" for language. Instead, patterns of early language style affect all aspects of pho-

nology, grammar, the lexicon, and perhaps social and cognitive domains outside of language-proper. Bates et al. (1985) characterized a set of pathways of variation in language development from first words to grammar in a sample of 27 children studied between the ages of 10 and 28 months. A wide variety of linguistic measures were included in this study, chosen to reflect aspects of variation proposed by earlier investigators. They include both receptive and expressive measures from structured testing, free speech, and parental report. Traditional measures like vocabulary totals and MLU were supplemented with detailed analyses of the flexibility and productivity of both lexical and grammatical structures. There was, then, a fair chance for dissociations between grammar and the lexicon to emerge. Nevertheless, there was no evidence for asynchrony between these hypothetical components of language. Instead, variation in language was distributed along such "horizontal modules" as rote versus analytic behavior, and comprehension versus production.

The picture is considerably more complex than the "two strand" approach in the individual difference literature might lead us to expect. At 10 to 13 months two factors emerged: an *analytic factor* (including comprehension vocabulary, referential style in both comprehension and production, and flexibility in the use of common object names) and a *rote factor* (based primarily on spontaneous speech with comprehension factored out). At 20 months, we now find three factors: *analyzed production* (including the range of meanings expressed in single and multiword speech), *rote production* (including a precocious overuse of inflections and function words, and a tendency to imitate a novel word taught in the laboratory), and a newly-emerging separate factor for *pure comprehension* (reflected in a variety of receptive language indices). By 28 months, we are back to two factors only: *production* (cutting across both lexical and grammatical measures) and *comprehension* (again including both lexical and grammatical tests). The cross-age correlations among these factors are also complex; comprehension at one age weaves back in to production at the next, and rote learning seems (at least temporarily) to lead nowhere at all after 20 months.

There are, then, at least *four* kinds of mechanisms involved in the dissociations observed in early language development: analytic learning (what MacWhinney [1978] refers to as "analysis" and what Peters [1983] refers to as "segmentation"), rote learning (what MacWhinney [1978] refers to as "amalgam acquisition" and Peters refers to as "extraction of forms for later analysis"), comprehension, and production. As far as we can determine from the Bates et al. study and related studies by other investigators, these four horizontal factors affect all the major components of language: word order, grammatical morphology, phonology, and lexical semantics. It is of course quite possible that other dissociations will emerge later in development, including a separation between grammar and the lexicon. Linguistic modules may have to be constructed and developed over time before they can attain a modular status (c.f.

Karmiloff-Smith, 1985). But they do not seem to play a role in the early stages—or, more precisely, they do not play any obvious role in creating the dissociations that underlie individual differences in language learning.

There is a rather inelegant clash between the detailed albeit contrasting models of language learning we find in the current volume, and the rather vague and old-fashioned notions of analytic versus rote and comprehension versus production. These characterizations of individual variation have more discriptive than explanatory value. How can existing theories be modified to take these findings into account? We suggest that all of the models discussed here (including the competition model) have so far ignored some essential aspects of information processing and learning: perception, memory, attention and motivation. To preserve the computer metaphor that motivates so many current models, we might think of these in terms of the "peripherals" that subserve the language learning device, or mechanisms that select, preprocess and transfer the data on which the learner will operate. Individual differences in children may derive from variation in the status of these data-transfer systems.

First, at the perceptual level, all of the models discussed so far assume that the sentence strings input to the learner are exactly the same for the child as they are for the observer. In our elucidation of the competition model, we showed that cue cost factors like "perceivability" can influence language learning. But we have treated "perceivability" as a property of the stimulus. Certainly this must be true, to a considerable extent. There are psychoacoustic realities that enter into a determination of which cues to meaning will be learned first. But these will undoubtedly interact with both "software" and "hardware" properties of the perceiver. Consider a finding in the adult psycholinguistics literature (Pollack and Pickett, 1964): when words uttered in a conversation between adults are taken out of context and presented in a list format, only 47% are recognizable to other adult listeners. In other words, complete perception of the elements in a sentence stimulus presupposes the lexical, grammatical and pragmatic knowledge necessary to disambiguate individual elements. But even in the most radical nativist theories, children cannot be assumed to have such knowledge at the beginning of language acquisition. The sentence strings input to the learner must surely be perceived—at least initially—as strings of incomprehensible noise punctuated by islands of familiarity. Among other things, there may be individual differences in the absolute acuity of the auditory system making it possible for some children to literally hear more of the stimulus prior to an analysis of its meaing. We know, for example, that "hyperacousis" is a characteristic of Williams Syndrome in children who are retarded but have surprisingly advanced language. Like Radar O'Reilly in *Mash,* these children can reportedly hear a helicopter coming long before anyone else nearby. Even in the normal range, production may move ahead of comprehension, and rote strings may outnumber analyzed strings in children who quite simply hear better.

By the same line of argument, the strings input to the learner presumably have to be held in some kind of memory buffer long enough for "on-line analysis" and/or storage to take place. We suggested earlier that the operation of cue validity in acquisition is tempered by the cue cost factor of "assignability"—once again, viewed as a property of the stimulus. The ability to hold elements in memory long enough to make an assignment and/or to store the form for later consideration, may also be a variable property of the learner herself. The claim that so-called pronominal/expressive children are particularly good at early prosody provides some support for this notion, suggesting that these children perceive and maintain the "intonational envelope" of a sentence that carries many of the otherwise incomprehensible phonetic units with it. Similarly, in some reports of retarded children with exceptionally good expressive language (e.g., Curtiss, Yamada & Kempler, 1981), the children seem to have exceptionally good auditory short term memory spans, at or above their age level and far ahead of their performance on cognitive tasks.

Attention is another domain that must be involved in determining the shape of language learning. The "inner searchlight" discussed by some investigators may focus on different parts of the linguistic signal with important results for the kinds of form/function correlations that will be noticed. We know from studies of eye movement monitoring that there are individual differences in the nature of the scanning process in visual pattern perception. Some analogue to scanning must play a role in language as well in order to determine which parts of the input will actually be learned, and when the learning will occur. As Bowerman (this volume) has pointed out, the models discussed here have been fairly agnostic about exactly when all this learning takes place. Does it happen "on line," as the stimulus comes in? Or does the child re-present stimuli to herself later on, changing her system in the process? The so-called referential/analytic child may actually be more attentive, spending more time focussing on the segmentation of utterances both inside and outside of the real-time process of communication. Such a difference in the distribution of attention could account for several facts about referential/nominal/analytic style:

1. referential style tends to be associated with precocity in language development overall;
2. referential children tend to produce more overgeneralizations in grammatical morphology (a result which would occur if analyses are being forced at an early point in development when the competition pool is still rather small);
3. referential children are more consistent in their apparent rule systems (or whatever the analogue to rules may be in parallel distributed processing systems) as though they were continually trying to "clean things up" and eliminate exceptions.

Finally, we may want to consider the role of motivation in the tendency for children to focus their attention on particular kinds of language, start up conversations on particular kinds of topics, and demand particular kinds of linguistic data from their conversational partners. We now have several lines of evidence to suggest that referential/analytic children play more with objects and use objects to interact with adults even during the prespeech phase (e.g., Nelson and Furrow, 1984; Goldfield and Snow, in press). Perhaps for this reason they also tend to elicit more object-oriented, nominal language from their partners. Hence the initial predisposition of a child to focus on the world of reference may set up a benign cycle of language-about-objects that feeds into an increasingly consistent language style. Similarly, expressive/morphological/rote styles of learning typically involve a great deal of immediate imitation of adult input. These children tend to be more sociable overall; they engage in a wider variety of social/regulatory speech acts, and seem to want to jump into the game of conversation without going through the tedious process of figuring out the linguistic pieces first. As several investigators have suggested, they learn "the tune before the words." Such an imitative style will insure that the child has a larger and more heterogeneous array of elements stored in the linguistic anteroom, awaiting further analysis. The nature of the "competition pool" will be very different for this child than it is for her more reflective and analytic counterpart.

In short, we cannot assume that the data input to the learner is the same for every child—even though it may look the same from our point of view. The learning systems described by the contributors to this volume—whether they are nativist or empiricist, deterministic or probabilistic—are necessarily embedded within a much more complex, multi-component system for selecting, shaping and delivering both the input and output of learning. This will of course make it more difficult to model the language learning process, since more components must be modelled. But if we can account in a principled way for variation in language learning, we can be much more comfortable in concluding that we have verged on something like the truth.

REFERENCES

Ammon, M. S., & Slobin, D. I. (1979). A cross-linguistic study of the processing of causative sentences. *Cognition, 7,* 3–17.

Bates, E. (1976). *Language and context: Studies in the acquisition of pragmatics.* New York, NY: Academic Press.

Bates, E., Bretherton, I., & Snyder, L. (in press). *From first words to grammar: Individual differences and dissociable mechanisms.* New York: Cambridge University Press.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A cross-linguistic study of the development of sentence interpretation strategies. *Child Development, 55,* 341–354.

Bates, E., MacWhinney, B., & Smith, S. (1983). Pragmatics and syntax in psycholinguistic research. In S. Felix & H. Wode (Eds.), *Child language at the crossroads.* Tubingen: Gunter Narr.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A. & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition, 11,* 245–299.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language.* New York, NY: Wiley.

Bever, T. (1975). Cerebral asymmetries in humans are due to the differentiation of two incompatible processes: holistic and analytic. *Annals of the New York Academy of Sciences, 263,* 251–262.

Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development, 40,* whole no. 2.

Braine, M. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development, 41,* Whole No. 1.

Bretherton, I., McNew, S. Snyder, L., & Bates, E. (1983). Individual differences at 20 months: Analytic and holistic strategies in language acquisition. *Journal of Child Language, 10,* 293–320.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Brunswik, E. (1956). *Perception and the representative design of psychology experiments.* Berkeley, CA: University of California Press.

Carroll, J. M. (1978). Sentence perception units and levels of syntactic structure. *Perception and Psychophysics, 23,* 506–514.

Chapman, R. S., & Kohn, L. L. (1978). Comprehension strategies in two- and three-year-olds: Animate agents or probable events? *Journal of Speech and Hearing Research, 21,* 746–761.

Chomsky, N. (1982). *Lectures on government and binding.* New York, NY: Foris.

Curtiss, S., Kempler, D., & Yamada, J. (1981). The relationship between language and cognition: Theoretical framework and research design. *UCLA Working Papers in Cognitive Linguistics, 3,* 1–60.

Devescovi, A., & Caselli, C. (1985). *Acquisition of subject-verb agreement by Italian children.* Manuscript, University of Rome.

Devescovi, A., & Taeschner, T. (1985). Strategie di uso del pronome in eta' scolare. Manuscript, University of Rome, Faculty of Psychology.

Furrow, D., & Nelson, K. (1984). Environmental correlates of individual differences in language acquisition. *Journal of Child Language, 11,* 523–534.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston, MA: Houghton Mifflin.

Gleitman, L., & Wanner, E. (1982). Language acquisition: the state of the state of the art. In E. Wanner and L. Gleitman (Ed.), *Language acquisition: the state of the art.* New York, NY: Cambridge University Press.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674–691.

Goldfield, B., & Snow, C. (in press). Individual differences in language acquisition. In J. Gleason (Ed.), *Language development.* Columbus, OH: Merrill.

Gvozdev, A. N. (1961). *Voprosy izucheniya detskoy rechi.* Moscow: Akademija Pedagogika Nauk RSFSR.

Hakuta, K. (1982). Interaction between particles and word order in the comprehension of simple sentences in Japanese children. *Developmental Psychology, 18,* 62–76.

Horgan, D. (1981). Rate of language acquisition and noun emphasis. *Journal of Psycholinguistic Research, 10,* 629–640.

Ito, T. (In press). Universal and particular in sentence processing. In B. MacWhinney and E. Bates (Ed.), *Cross-linguistic studies of sentence processing.* New York, NY: Cambridge University Press.

Johnston, J., & Slobin, D. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language, 6,* 529–545.

Kail, M., & Combier, C. (1983). Role des indices pragmatiques, lexico-semantiques et syntaxiques

..dans la comprehension de phrases simples dans une perspective genetique et interlangue. Master's degree, University Rene Descartes.

Kail, M., & Hernandez-Charvillat, A. (1983). Role des indices syntaxiques, semantiques et pragmatiques dans les strategies de comprehension chez les adultes francais. Master's degree, University Rene Descartes.

Kempler, D. (1980). Variation in language acquisition, UCLA Working Papers in Cognitive Linguistics. UCLA Linguistics Department.

Li, C., & Thompson, S. (1976). Subject and topic: a new typology of language. In C. Li (Ed.), *Subject and topic.* New York, NY: Academic Press, Inc.

Lightfoot, D. (1982). *The language lottery.* Cambridge, MA: MIT Press.

MacWhinney, B. (1975). Rules, rote, and analogy in morphological formations by Hungarian children. *Journal of Child Language, 2,* 65–77.

MacWhinney, B. (1975). Pragmatic patterns in child syntax. *Stanford Papers And Reports on Child Language Development, 10,* 153–165.

MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development, 43,* Whole no. 1.

MacWhinney, B. (1984). Grammatical devices for sharing points. In Schiefelbusch, R., & Pickar, J. (Ed.), *The acquisition of communicative competence.* Baltimore, MD: University Park Press.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior, 23,* 127–150.

MacWhinney, B., Pleh, Cs., & Bates, E. (1985). The development of sentence interpretation in Hungarian. *Cognitive Psychology, 17,* 178–209.

MacWhinney, B., & Price, D. (1980). The development of the comprehension of topic-comment marking. In D. Ingram, C. C. Peng, & P. Dale (Eds.), *Proceedings of the First International Congress for the Study of Child Language.* Lanham: University Press of America.

Maratsos, M. (1982). The child's construction of grammatical categories. In Wanner, E., & Gleitman, L. (Ed.), *Language acquisition: The state of the art.* New York, NY: Cambridge University Press.

Maratsos, M. (1983). Some current issues in the study of the acquisition of grammar. In Flavell, J., & Markman, E. (Ed.), *Handbook of Child Psychology: Vol. 3.* New York, NY: Wiley.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88,* 375–408.

McClelland, J., & Rumelhart, D. (1986). A PDP Model of the Acquisition of Morphology. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, N.J.: Lawrence Erlbaum.

Nelson, K. (1973). Structure and strategy in learning how to talk. *Monographs of the Society for Research in Child Development, 38,* Whole Nos. 1–2.

Nelson, K. (1981). Individual differences in language development: Implications for development and language. *Developmental Psychology, 17,* 170–187.

Peters, A. (1983). *The units of language acquisition.* New York, NY: Cambridge University Press.

Pinker, S. (1982). A theory of the acquisition of lexical-interpretive grammars. In J. Bresnan (Ed.), *The mental representation of grammatical relations.* Cambridge, MA: MIT Press.

Plèh, Csaba. (1981). Kŭlönbözö szörendü mondatok èrtelmezèse ès a dichotikus hallàsi aszimmetriàk 3-6 èves gyermekeknèl. *Pszichològia, 1,* 365–393.

Pleh, Cs. (1981). The role of word order in the sentence interpretation of Hungarian children. *Folia Linguistica, 15,* 331–343.

Pollack, I., & Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior, 3,* 79–84.

Posner, M., & Keele, S. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Radulovic, L. (1975). *Acquisition of language: Studies of Dubrovnik children.* Doctoral dissertation, University of California at Berkeley.

Schachter, P. (1976). The subject in Philippine languages: Topic, actor, actor-topic, or none of the above. In C. Li (Ed.), *Subject and topic*. New York, NY: Academic Press, Inc.

Sinclair, H., & Bronckart, J. (1972). SVO—a linguistic universal?: A study in developmental psycholinguistics. *Journal of Experimental Child Psychology, 14*, 329–348.

Slobin, D. (1985). *The cross-cultural study of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Slobin, D., & Bever, T. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition, 12*, 229–265.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development*. New York, NY: Holt, Rinehart, & Winston.

Smith, S., & Mimica, I. (1984). Agrammatism in a case-inflected language. *Brain and Language, 21*, 274–290.

Snyder, Lynn S., Bates, Elizabeth, & Bretherton, Inge. (1981). Content and context in early lexical development. *Journal of Child Language, 8*, 565–582.

Strohner, H., & Nelson, K. E. (1974). The young child's development of sentence comprehension: Influence of event probability, nonverbal context, syntactic form, and their strategies. *Child Development, 45*, 567–576.

Thibadeau, R., Just, M., & Carpenter, P. (1982). A model of the time course and content of reading. *Cognitive Science, 6*, 157–203.

Vihman, M. (1985). Individual differences in babbling and early speech: Predicting to age 3. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech*. Basingstroke, Hampshire: MacMillan.

Vihman, M., Ferguson, C., & Elbert, M. (in press). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*.

Warren, R., & Warren, N. (1976). Dual semantic encoding of homographs and homophones embedded in context. *Memory and Cognition, 4*, 586–592.

Weist, R. & Konieczna, E. (1985). Affix processing strategies and linguistic systems. *Journal of Child Language, 12*, 27–36.

Xiao-chun, M. (1981). Word order and semantic strategies in Chinese sentence comprehension. *International Journal of Psycholinguistics, 8*, 23–33.