
1-2001

Continuity and Modularity in Language Acquisition and Research

Jill de Villiers

Smith College, jdevilli@smith.edu

Follow this and additional works at: https://scholarworks.smith.edu/phi_facpubs



Part of the [Linguistics Commons](#), [Philosophy Commons](#), and the [Psychology Commons](#)

Recommended Citation

de Villiers, Jill, "Continuity and Modularity in Language Acquisition and Research" (2001). Philosophy: Faculty Publications, Smith College, Northampton, MA.
https://scholarworks.smith.edu/phi_facpubs/47

This Conference Proceeding has been accepted for inclusion in Philosophy: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu

Continuity and Modularity in Language Acquisition and Research

Jill G. de Villiers

Smith College

ARLA 3nd Draft March 2001

1.0. Overview

The paradigm of generative approaches to language acquisition is the focus of critical attention here. Although much ink has been spent on the comparative merits of different approaches, this can detract from attention to the details within a paradigm.

Thus what is not part of this paper is any attempt to compare competing paradigms in language acquisition research. The task of reviewing twenty years of intense research is too grandiose: instead, four representative lines of empirical work with broad theoretical significance and a substantial cross-linguistic base are chosen to illustrate how debates within the paradigm are structured. Theoretical and methodological commitments of researchers can be seen to vary considerably, and most importantly, to interact. The goal is a selective map to reveal the structure of arguments, and to ask whether we are making progress.

The paper is structured as follows. In Section 1.1 the two basic principles that provide the title to this paper are introduced, together with a sketch of the reasons why the four empirical examples are chosen. In Section 1.2 a more complete history of the principles is given, after which in Section 2, the general character of the arguments used to defend them is provided. Finally, in Section 3, each empirical line of work is discussed in the light of the principles and the arguments that maintain them.

1.1 Continuity and Modularity

The generative approach to language acquisition enshrines two principles, namely, Continuity and Modularity. These principles are considered paramount. Research that is in support of, or maintains, these principles is rated highly by the community of researchers. Research that produces results that call them into question is marginalized, or a flurry of research activity begins with the goal of demonstrating that the results are artifactual. That is how paradigms in science are maintained: it might seem overly conventional, but the advantages of conventionalism lie in the lack of distraction and the subsequent attention to detail that paradigms encourage (Kuhn, 1963).

The concept of **Continuity** was coined first by Macnamara (1982) and elaborated in Pinker (1984), it has since been emphasized in most work in the generative tradition (Lust, 1999; Wexler, 1999). This quote is representative, referring to the Strong Continuity hypothesis:

“UG (where this term refers to the “principles and parameters” which provide the true content of UG) is a model of the Initial State: it is thus available to the child from the

beginning. The initial state is taken to refer to the onset of first language acquisition, even before experience. UG remains continuously available throughout the time course of first language acquisition. UG does not itself change during this time course”.

(Lust, 1999, p. 118)

i.e. that the null hypothesis should be that *a child's grammar has all of the features and properties of adult grammar.*

Fodor (1981) was the originator of the general notion of **Modularity** of cognition, and he put forward several important proposals about the properties that modules should have: specialized kinds of input, fast, mandatory, not accessible to introspection, and encapsulated, i.e. *impervious to the influence* of general knowledge or cognitive processes. He used language phenomena as two of his best examples: syntax and speech were each held to have many of the classic properties of modules.

In what follows, two cases are chosen from the child's earliest grammars to illustrate the kinds of reasoning that these fundamental principles have occasioned over the past decades. What the two debates have in common is a focus on the nature of missing elements in early grammar, namely, why is the child's output reduced? Why is early English “telegraphic” in nature (Brown, 1973)? The obvious (and perhaps sensible!) responses are two-fold:

- i. The child doesn't yet know language, so he has to start simply, building up grammar as he goes.
- ii. The child knows language, but the child has cognitive limitations in either expressing or comprehending sentences: limited memory, limited attention, limited world knowledge.

But note: i) threatens Continuity, and ii) threatens Modularity.

The first case to be discussed (Section 3.1) reflects this larger issue of reduced output: in particular, is the full skeleton of adult grammar, X' theory, really there in the child but mostly empty? Or, does the child's grammar consist in some subset of possible categories, say only lexical phrases?

The second case (Section 3.2) concerns a particular form of reduction, namely missing subjects in child speech. Why are subjects missing and not objects? Both cases

allow full exploration of the styles of reasoning identified below.

Two further cases are then chosen to illustrate contemporary discussion of the issues in later grammar. Both reflect not the analysis of spontaneous speech of children but rather experimental results, typically in comprehension experiments. In each case, the child is observed to do something principled and adult-like, and something else paradoxical and non-adult-like. Once again, the reasonable solutions fall into two classes:

The child doesn't yet know all the grammar involved.

The child knows the grammar but gets misled by conditions in the experiment.

And once again, the dilemma is that i) violates Continuity and ii) violates Modularity.

In Section 3.3 I discuss the Principle B binding "error", one of the most exhaustively studied phenomena in the past twenty years of research. It provides an interesting window onto the struggle to maintain both principles, in the light of changing linguistic theories. In Section 3.4 the second example, of Wh-barriers, leads by accident into an interesting set of questions about Modularity per se.

1.2 Some recent history of the principles.

Chomsky appealed to two forms of adequacy of linguistic description: descriptive and explanatory. By explanatory adequacy he means:

“ a theory of language must show how each particular language can be derived from a uniform initial state under the ‘boundary conditions’ set by experience”

(Chomsky 2000), p. 7).

The appeal to continuity of grammars between child and adult is an attempt to satisfy explanatory adequacy in recognition of the learnability problems introduced by positing discontinuity. If there is change between child and adult grammar, then how are those changes accomplished? Given the other fundamental assumptions of the paradigm, that there is no negative evidence available to the child, that the principles proposed as universals are highly abstract, and that the language input is impoverished in important ways, learning theories are faced with an impossible challenge. So generative approaches to language acquisition generally place explanatory adequacy higher in importance, at the expense of positing descriptions that are sometimes far removed from the surface forms of language.

The proposal of Continuity was a constraint surprising to many who worked on child language within an older tradition, in which descriptive adequacy was highly ranked. Anyone working on early child language was struck first by the apparent discrepancies between what children said or understood compared to what adults knew: the goal of many descriptive studies was to detail the changing grammars of the young child as they moved through “stages” such as telegraphic speech. But the proposal of Continuity is that positing grammars based on different principles is to be a last resort: the null hypothesis is that the child knew everything that the adult knew about grammatical principles. In some work, the notion of Continuity has been extended beyond principles to structures, e.g. X’ theory. Some have extended it to processing, still others to pragmatic principles. It is only a matter of time before it is extended to the lexicon too: Chomsky (2000) has recently emphasized how deep the young child’s knowledge of word meaning is, compared to what the input can possibly supply (see also Bloom, 1999). So Continuity is a basic thesis, and the consequences for dispensing with it are serious ones given the learnability assumptions.

However, Borer and Wexler (1987) wrote, “Linguistic explanations of ordering in development, based on the continuity hypothesis, involve the assumption of some kind of ordering in linguistic theory (extrinsic or intrinsic). And we have suggested there is no *linguistic* motivation for the ordering....we then conclude that linguistic theory might be preferable without the continuity hypothesis, if the alternative allowed the non-principled assumptions to be removed from linguistic theory” (p. 126).

As to Modularity, the idea of a specialized language organ, a language module, is emphasized in many of Chomsky’s writings. However, in recent work he writes

“The faculty of language is embedded within the broader architecture of Mind/brain. It interacts with other systems, which impose conditions that language must satisfy if it is to be usable at all”

(Chomsky 2000, p. 9).

These legibility conditions imposed by the articulatory and cognitive systems are held to constrain the forms that language takes. By this he means constraints on the basic design of languages, not influences in the course of speech nor in ontogeny.

Crain and Thornton (1998) most recently adopt the Modularity Thesis as a fundamental tenet of language acquisition research, as do Lillo-Martin (1999), Crain and Lillo-Martin (1999) and Crain and Wexler (1999). Crain and Thornton make the case as follows:

“The construction of syntactic and semantic representations of sentences is not influenced by general cognitive mechanisms...Only the output from the language faculty makes contact with real-world knowledge....According to the Modularity Matching model, all of the linguistic abilities of a child are the same as an adult’s. Not only do we assume that children have access to Universal Grammar, but we also make the more controversial assumption that children are equivalent to adults in the mechanisms they use to process language; that is, they have access to a universal parser.”

(Crain & Thornton 1998, p. 30).

Crain and Wexler (1999) also assume that language is a module separate from other cognitive systems in the mind. They further assume that the human language processing system is modular, that is, the principles of phonology, syntax and semantics are all autonomous subcomponents within the larger language module. Most strikingly, they also assume as a null hypothesis that the child’s processing modules are each exactly like those of the adult. Crain and Thornton (1998) caution that to do otherwise is to open a “Pandora’s box” of alternatives.

Now it is interesting that these latest proposals about modularity in child language take a stricter view of the notion of modular language processing than has been taken in the past, in at least two ways.

- i. In adult sentence processing literature, modularity in the sense of encapsulated processing, was held to apply to *on-line processing*. But *decisions* based on that parsing, say choosing a matching picture, were then held to be a product of both the linguistic parsing and other factors, such as world knowledge or picture biases, that apply at the decision point. So tasks such as choosing a picture to match a sentence may involve parsing the sentence in a modular fashion, but that is then only one contributing factor to the subject’s eventual choice. With no

articulated theory of these other factors and how they interact, researchers moved instead to the methodology of on-line parsing in assessing adult linguistic competence, a recommendation also made for studying child language by McKee, Nichol and McDaniel (1993).

- ii. Pragmatics was once held to be world knowledge, outside of the language module, but in Crain and Thornton (1998) pragmatics is part of the language module itself, the part that then connects to real world knowledge (p 31, Fig 4.1) In Wexler (1999) at least part of pragmatics concerns the interface module between syntax and cognitive systems. The boundary, that is, which aspects of pragmatics are inside and which are outside, deserves some further clarification (see also Schaeffer, 2000).

Researchers in the generative tradition often argue that in contrast to the principled, theory-driven nature of linguistic explanation, psychologists know next to nothing principled about the rest of cognition, namely attention, working memory, set effects, biases and all the other “performance factors” that might interfere with the display of competence. But of course, researchers outside of the tradition can argue that psychology knows a good deal about such cognitive factors, and virtually nothing certain about linguistics! For example, in a recent review of MacWhinney (1999) by Sabbagh and Gelman (2000) say , “The reliance on domain-general mechanisms challenges researchers to consider known cognitive constants before appealing to *ad hoc* rules in accounting for a wide variety of linguistic behaviors”. So *ad hoc* is in the eye of the beholder.

2.0 The arguments used to uphold the principles

Why, then, does child language not converge instantaneously on the adult model? That is, why does language acquisition take so long? Despite the strong continuity claims, it is evident to everyone that children speak and understand quite differently than adults, at least for the first couple of years. Furthermore, their utterances often seem to be driven by radically different grammars than the adult model to which they are exposed. How can these observations be reconciled with strong continuity? The solutions to this

dilemma fall into three classes discussed next, and researchers in the field divide in terms of how seriously they take each one.

2.1 Performance considerations

Consider the first possibility, that performance is not always a reliable index of competence. In early child language research, the main source of evidence was corpora of child speech- initially not even video-taped but audio-taped, with contextual notes scribbled by the researcher (but never sufficient to the task). Since the advent of the CHILDES system (MacWhinney & Snow, 1985) , many more researchers have access to transcripts of child speech across languages, and the distance of the researcher from the context is increasing, at least until video accompaniments can be more easily accessed and stored. The ambiguity of spontaneous speech has long been a source of contention in language research, but asking young children for grammatical intuitions or doing on-line parsing studies with them is still a marginal alternative, suitable mostly for older children. Thus the main alternatives for assessing grammar in adults are not available to the child language researcher. A huge amount of attention has been paid to experimental alternatives to spontaneous speech, and significant innovations have arisen. Most promising are preferential looking techniques with infants (Golinkoff, Hirsch-Pasek, Cauley & Gordon, 1987; Hohle & Weissenborn, 2000) , elicited production procedures (Crain, 1991), stories with questions (de Villiers & Roeper, 1995), and the Truth Value Judgement task (Crain & Thornton, 1998). The issue of methodology is at the heart of proposals by Crain and Thornton (1998) and Lust, Flynn, Foley and Chien (1999) that researchers must take methodology seriously as a source of artifact in child language research, though their solutions differ. They agree, however, that many of the findings that young children's grammars may depart in significant ways from adult grammar could be artifacts of the procedures used to test them. One of the first demonstrations of this came from Hamburger and Crain (1982), arguing that the usual comprehension procedures for testing relative clause knowledge in children were flawed in their pragmatic/semantic properties. That is, a child would be typically given a sentence containing a relative clause to act out:

- 1) The dog that chased the donkey touched the pig

and the experiment would demand that the child enact both clauses to get it right, even though in the normal use of the relative clause the event in the relative is presupposed, it is used to identify which dog. When this demand is dropped and the situation made to match the normal pragmatics, children performed much better and more in keeping with the spontaneous speech evidence of rare but well-formed relatives. The basic theme has been repeated in much of Crain's work (Crain, Thornton, Boster, Conway, Lillo-Martin, & Woodams, (1996); Crain & Thornton, 1998): if you are going to test children with experimental materials, then make sure the pragmatics are felicitous for the structure in question.

In strong contrast to the Modularity matching model, and serving as its impetus, are the proposals of Competing Factors models (Bates & MacWhinney, 1989) and the Coalition Model (Hirsch-Pasek & Golinkoff, 1997). Crain and Wexler (1999) point out that in some such models, grammatical principles are not even included among the possible contributors to behavior. In the Coalition Model (Hirsch-Pasek & Golinkoff, 1997), the child goes through stages of differential bias to one set of cues versus another. At first, the cues are primarily prosodic, then semantic, then finally syntactic cues predominate. At each stage, however, the child is discovering language in the input. Even within the generative tradition, it is generally assumed that grammatical principles are one such factor, but that other factors such as attention, bias, memory load etc. might mask the contribution of those principles to performance. Hence at best one might expect a kind of rough statistical conformity of behavior to that predicted by the grammar alone.

Crain and Wexler reject such a model for approaching child language research. They argue that if the conditions are arranged correctly, namely "if children understand the experimental task and are cooperative" and "if children can parse the sentence and relate it to the context" (p. 391) they should be perfect or almost so, in their treatment of the grammatical principle involved. However, they also allow that the non-linguistic demands of an experimental task can affect the outcome, for example if the child has to execute a complex cognitive plan (cf. the critique of the work on relative clauses above). But Crain and Wexler argue that even when such demands result in consistent non-adult responses, it should not be concluded that the child has a different grammar. Instead, the extraneous sources of demand must be reduced, and the experiment re-done. They also

admit that children might make errors if they cannot assign a grammatical interpretation to a sentence that is “too long or too complex” (p. 393), and might be driven to invoke factors outside of the grammar to solve the task. In addition, Crain and Thornton (1998) claim that children may be subject to greater uncertainty than adults when they are put in unnatural experimental situations that are infelicitous pragmatically compared to normal language use.

Given this list of potential performance “excuses”, then how are the competing factors models and the Modularity Matching model distinct? Theoretically, the grammar is considered primary on the latter view, not just one of a set of factors or even the factor weighted more heavily than the others. But the Modularity Matching model recognizes also that all the other variables have to be “controlled” before the result can be taken to be a reflection of grammar. The danger is that nothing less than adult performance will be accepted as evidence. If the result is just statistically in favor, not absolute (with some allowance for “noise”), then the experiment has not been done correctly. This is presumably judged by the outcome of pilot work, and careful thinking about the problem. In addition, the claim of strong modularity is that adults should be prone to make the same mistakes on the task if the demands of the task are extraneous, since the adult’s grammatical knowledge is not in question (though there are in fact multiple proposed differences between child and adult, as seen above). And the null hypothesis is that the child’s capacities within all the linguistic modules at least, are equivalent to the adult’s. Empirically, this is a hard line to take, because it entails the pursuit of perfection wherein anything less is taken as due to experimental mistakes.

Paradoxically, then, the very thesis that the child’s grammar is strictly modular, has the potential to lead to the most extreme focus on extraneous sources of performance error. Crain and Thornton take the approach that experimental work in the field should heed their call to use only two methodologies that are relatively impervious to performance error, which in their judgement are elicited production, and the truth value judgement task. In elicited production, a scenario is set up to encourage production of what might otherwise be a rare form, say a relative clause or an embedded question. The child is led to say the critical sentence through engagement in a conversation with the experimenter and a puppet, for example, and is told that the puppet either needs to know

something, or will only speak to the child so instructions have to be conveyed through the child. In the truth-value judgement task, a scenario is enacted in front of the child using props, then a description of the events is provided by a puppet who is just learning to talk, so the child must judge whether what the puppet says is appropriate or not. However, Lust et al (1999) consider the truth value judgement task to be an example of a “reduced behavior task”, in which the binary choice of a yes or no answer masks the processes that led to the choice, and can therefore mislead a researcher. Similar criticisms are launched against the preferential looking task in infants, another reduced behavior task. Elicited production tasks do not have the same objection, but they are not always possible to design without leading the child, and the Truth Value Judgment Task has become the method with the widest applicability, and some stunning demonstrations of success (Crain & Thornton, 1998).

The message of the Modularity Matching model is best summarized as follows: do not conduct a loosely controlled experiment and then explain the child’s behavior in terms of a combination of competence and performance influences. Instead, get the experiment right, so the child’s behavior is not just statistically in favor of one hypothesis or the other, but conclusively so, with maybe 10% error allowed.

2.2 Parametric variation

The second option is also widely accepted: the child’s grammar differs from the adult’s grammar not because of lack of UG principles, but because the child has not set the appropriate parameters for his or her particular language. In other words, the child’s grammar falls under some other possible instantiation of UG. This permits Continuity to be maintained, but now learnability problems potentially arise for parameter setting (Nishigauchi and Roeper, 1987).

An important contribution of the Principles and Parameters approach is that the variation across languages be systematized into a small set of parameters (estimates are 30- 40 in Clark, 1992) such as Head direction in phrases, wh-movement, V-2 movement and so forth. Instead of accumulating evidence for how the language behaved, the child was essentially looking for evidence that might set a relatively small number of binary options, switches. Ideally, there might be linkages among the parameters, so some might

get set for “free” (Wexler & Manzini, 1987; Nishigauchi and Roeper, 1987). Or there might be unambiguous evidence, “triggers”, guaranteed to be available to the language learner whatever his or her circumstances of input (Fodor, 1998).

What is at issue here is at the heart of the parameter-setting idea. Parameters were proposed to delimit the set of variations on Universal Grammar exemplified across languages: if the child had a set of switches for different properties such as Head Direction, Null Subject, Wh-movement etc., then in essence the child would “know what to look for” in the evidence: the theory of grammar is built in, and not constructed on the basis of hypothesis testing. But the theory depends critically on having evidence available - a Unique Trigger, in Roeper and de Villiers (1992) - that unambiguously sets the parameters. If that data just consist of cumulative evidence that the language does it this way rather than that way, then why not revert to hypothesis testing (Cowie, 1999)? In defense of parameter setting, it is because the hypotheses are severely constrained in advance, rather than emergent. This point will be elaborated below, with a new solution recently proposed.

Gibson and Wexler (1994) proposed a Triggering learning algorithm whereby the learner changes the value of a parameter whenever the current grammar cannot analyze an incoming sentence. The problems that arise have to do with the existence of so-called “local maxima”, that is, grammars that are not the target grammar, but resemble it sufficiently to be chosen, with no possible escape. That is, every departure from the local maximum suggests the choice is worse, but it is not the “peak”. Mathematical modeling with say, 12-parameter spaces, suggests the existence of many such local maxima and hence a serious problem for the algorithm (Kohl, 1999, cited in Yang, 2000).

However, to the extent that the evidence is not unambiguous, or that triggers are not immediately available, then there are several possibilities under this approach:

The child might set a parameter to some default setting.

The child may set a parameter to the wrong setting and have to retreat.

The child may not commit to any parameter setting, but have multiple grammars.

These alternatives have all found advocates in contemporary research over the last fifteen years, and sufficient uncertainty and indeterminacy has promoted some to advocate

instead very early parameter setting (Wexler, 1998), thus back to the strongest of competence continuity theories. Others (Roeper, 1998; Roeper, 1996; Nishigauchi and Roeper, 1987) have maintained that parameter-setting involves the accumulation of features on categorial heads (Chomsky, 1995), hence a necessarily slow process more akin to lexical learning. So can both very early and late parameter setting co-exist, and refer to different parameters?

2.3 Maturation

The instantaneous hypothesis (Lust, 1999) is a claim that the principles of UG provide a model of the initial state of the child. There is no change over development in UG: it is there prior to experience and it does not undergo maturation. Borer and Wexler (1987) made the first specific proposal for maturational change in UG to provide an account for children's difficulty with verbal passives. They argued that A-chains were unavailable in the child's grammar until later in childhood (4 or 5 years), and that the model they proposed entailed that

“in contradiction to the continuity hypothesis...the principles are not available at a certain stage of a child's development, and they are available at a later stage” (p. 124, italics added)

Borer and Wexler were concerned with the triggering problem in development: why, if data were available in the input to set a parameter, did language learning take so long? They reject the possibilities that learning of UG principles occurs, or that the input is structured so as to delay acquisition.

Lust (1999) attacks the specific grammatical maturational thesis on the grounds that it is unprincipled. Biology determines not only UG, but also the ordering and the stages of growth that arise on the route to a full adult grammar, a kind of stage-wise unfolding. But Lust argues that the description of these stages follows from nothing in Linguistic theory, rather, the proposal is “just” empirically motivated by the behaviors that are seen in children. Furthermore, since UG is not the initial state but the end-state that arises out of this biological program, then “it robs the linguistic theory of UG of its essential scientific force... because UG was proposed in its status as a model of the initial state” (p. 125). In particular, the logical problem arises that UG can not be used to

validate the empirical generalizations made about development, since it is divorced from the initial state. This is a curious twist on the quote above from Borer and Wexler (1987), who also argued that if linguistic theory had little principled to say about the ordering of development, that it might be better to forego Continuity!

In a more recent formulation, Wexler (1999) reiterates that there are two components to the development of language in the child:

a genetic program guiding the growth of grammar

a learning component for the language-specific aspects of grammar.

However, his claim in that paper is that all children's grammatical representations must be available in UG, i.e. maintaining Continuity, unlike the formulation in Borer and Wexler (1987). Maturation is invoked to explain why the child's grammar is not the adult's grammar, but not to explain representations not found in UG. What the maturation thesis does allow is that certain representations – like A-chains - are missing altogether from the child's grammar.

Wexler (1999) puts forward the proposal that what in fact might mature are the interface properties of language. In Minimalism (Chomsky, 1995) there are two interfaces: the phonological interface between phonology and phonetic form (PF), and the conceptual interface between logical form (LF) and the conceptual interpretive system. Wexler hypothesizes that some of the delays in child language might result from a lack of co-ordination between the computational/syntactic aspects of language, and the conceptual system.

Lust (1999) acknowledges that the notion of UG-constrained maturation in the later proposals by Wexler (Chien & Wexler, 1990; Wexler, 1999) is much more compatible with her own position, in that something else matures instead of UG. Lust argues that the Strong Continuity Hypothesis can be maintained without recourse to modifications in UG across time. Her own work explains developmental delays on the basis of the interaction of modules that must be integrated in a language-particular manner, for example parametrically determined head direction and the binding principles (Lust, 1986), or the syntax integrated with aspects of the lexicon (Mazuka & Lust, 1994).

2.4 Summary

Performance considerations, parametric variation, and maturation are the three common forms of explanation for departures from adult grammar seen in the grammar of young children. All represent attempts to preserve the principle of Continuity, though they do so at a further cost. Invoking **performance considerations**, an old standby for preserving Continuity of grammar, is now under attack because it runs counter to Modularity considerations. Invoking **parametric variation** is an attractive alternative that runs up against potentially serious learnability questions. **Maturation** raises the hard question: maturation of what? If grammar, then Continuity becomes questionable; if something else that affects grammar, then Modularity becomes questionable. These are considerations worth weighing as the particular examples are discussed.

3.0 Empirical examples

We turn now to the specific examples of empirical traditions that come into contact with these arguments.

3.1 Functional categories

According to X' theory (Jackendoff, 1977; Chomsky, 1995), all the phrasal categories had the same basic shape, with a head that is the key constituent: N for NP, V for VP and so forth. The X' structures include not only lexically headed phrases such as noun phrase (NP), verb phrase (VP), and adjectival phrase (Adj P), but also “functional” phrases IP (inflectional phrase) and CP (complement phrase) and DP (determiner phrase). These functional categories are headed by functional components, I (for inflection), C (for complementizer) and D (for determiner). The heads of lexical categories have categorial features, grammatical features (phi features checked during derivations), phonological properties, and features determining semantic (s)-selection and categorial (c)-selection. Functional heads have also feature structure, but they do not participate in theta-marking.

The top of the clause is typically taken to be a CP, that in a subordinate clause can be headed by a complementizer. In a root clause the CP becomes the landing site for movements e.g. of wh-words (to spec-CP), auxiliaries in questions (I to C movement) and verbs in V-2 languages (V to C).

The complement of C is then the IP category, held responsible for tense and for assigning the subject case in the sentence, and I (or INFL) is the key constituent of that phrase. The Spec of IP contains the subject. In English INFL can contain a modal (e.g. “will”) or a form of tense. Verbs are either nonfinite (e.g., the infinitive in English such as “to play”), or finite (e.g., “playing”, “played”, “plays”). Nonfinite forms can only occur in subordinate clauses e.g.

2) I wanted him to play.

3) He saw the girl play.

but all main clauses require a tensed, or finite, form. The tense of the sentence is determined by the component I. The nominative case on a subject is dictated by the I component: If the clause is tensed, then the case of the subject is nominative:

4) She heard the bell.

5) He was wondering if she wanted a snack.

but if the clause has no tense, then the subject is accusative:

6) She saw him weep.

7) I wanted them to leave.

In subsequent formulations (Chomsky, 1992), movement of elements is motivated by checking: the verb moves to I to check tense and/or agreement, the subject moves to I to check case. Other proposals refined the INFL components into two nodes: AGR (agreement) and T (Tense) (Pollock, 1989).

The DP is the site for determiners such as definite articles (Speas, 1990). Longobardi (1994) argues that the role of D is to allow the noun to be interpreted as an argument. Without D, nouns are predicative, naming kinds, not specifics (Stowell, 1991; see also Chierchia, 1998).

Several theorists accounted for the obviously restricted output of children’s early sentences by appealing to the notion that the functional categories were absent from the first grammars. Radford (1988, 1990a) was among the first to consider the implications of X’ syntax and functional categories for children’s early two-word sentences. Radford’s claim was that children’s grammars lacked the functional categories at the beginning. As a result, their earliest sentences were held to be constructed only of lexical-thematic

categories, such as NPs and VPs, on analogy to a Small Clause in adult grammar (Stowell, 1981). Lebeaux (1988) also argued that functional categories might be absent in children's initial grammars of English, as did Guilfoyle and Noonan (1988) and Platzack (1990) for early Swedish. The basic idea was also extended as a description of the impoverished English used by oral deaf children (de Villiers, de Villiers & Hoban, 1994).

For example, children's first sentences frequently lack auxiliaries or tense markers:

8) Me go home.

9) Teddy have cake.

They frequently lack the infinitival "to" that appears in adult (embedded) sentences that do not have a tense marker:

10) Adult: I want this to fit here.

11) Child: Want this fit here.

Children's early sentences seem to lack nominative case on the subject of the sentence:

12) Me want that book.

Radford thus made a reasonable case that children's early grammars lack the IP category, and similar arguments can be made for the CP category (Radford, 1990a), and the DP category (Radford, 1990b). Other theorists made the case that the full set of functional categories might be initially reduced, e.g. a single undifferentiated category that becomes IP and CP (e.g. a Finiteness phrase, Clahsen, 1990; Meisel & Muller, 1990). This idea of categories without strict labels became more palatable within Minimalism (Chomsky, 1995).

3.1.1 Performance considerations

Radford's (1990a,b, 1994) reports were devoid of statistical analysis: he simply provided several examples, often from different children, to illustrate each point. The conclusion about absent functional categories was soon called into question by other work, some of which examined the English data more quantitatively, and some of which called upon evidence from languages other than English. The major objection was that the proposal violated Continuity: children's grammars that consist only of lexical

categories represent options not found among the world's languages.

Within a short time, evidence began to accrue that the absence of functional elements in production should not necessarily be taken as evidence that the corresponding functional category is absent. To quote Poeppel and Wexler (1993), "absence of evidence for some category does not entail evidence of its absence". Several researchers noted syntactic effects of a functional category despite missing morphology. For example, for INFL, Pierce (1992) demonstrated that children move finite verbs to I across negation in French even when they are using bare forms (see below). For C, young German children move finite verbs to C in root clauses despite having in their speech no evidence of complementizers occupying that position (Poeppel and Wexler, 1993). For D, young Welsh speakers restrict their possible orderings of adjectives, numbers and possessives, suggesting a richly structured determiner phrase, before they produce definite articles (Aldridge, Borsley, Clack, Creunant & Jones, 1997). This kind of indirect evidence of the presence of empty functional categories has been influential in swinging opinion from the no-functional-categories idea, but what explanation is invoked for the absent inflections themselves?

The slow development of inflectional paradigms, even if not uniformly true, is not well-explained within generative models (neither will it yield to Connectionism: Pinker & Prince, 1988) and explanations tend to appeal to input frequency, or to the complexity or peripherality of the paradigms. For instance, Pierce (1992) explains the lack of inflectional marking in early English by saying for instance, that the morphology of 3rd person 's' marking habitual activity is impoverished in the input. But frequency has never worked as a solution (Brown, 1973). Perhaps it is significant that in English the inflectional system is peripheral (Hyams, 1986) compared to languages like Spanish and Italian, where more marking is seen early (see also Verrips and Weissenborn (1992) on German). But Poeppel and Wexler (1993) and Wexler (1996) also argue that matrix infinitives provide a grammatical alternative for the child, the so-called "optional infinitives" stage. They argue against processing or output omission models as explanations for the reduced form, so the logical alternative, given full competence and full functional categories, has to be optionality of output as a grammatical option for children.

Poeppel and Wexler (1993) offer no definite explanation for missing C, except that evidence of subordination is questionable in early German and hence one would not expect complementizers to show up. But they do hint at performance limits: “Whatever the reason for the absence of embedded clauses may be (say e.g. processing limitations, memory limitations), to argue that a functional category is not realized in the grammar on the basis that an expected constituent cannot be documented in the data is unconvincing and logically flawed”.

It would seem especially important that researchers bolster ambiguous production data with data from comprehension of the absent morphology (de Villiers, 1992; McKee, 1994). One of the first pieces of evidence that children in the initial periods of acquisition may understand something of the meanings carried by morphology before they master it in production comes from a study by Katz, Baker, and Macnamara (1974). The study is justifiably much-cited, because it promised hope of convergent methodology. The authors asked whether children at 17 months of age, could respond to certain meanings conveyed by articles, even though they did not yet use articles in their own productions. The articles carried semantic differences between a proper noun, i.e. a name, and a common noun. They showed children a new toy, either a doll or a block, and said,

13) This is Zav. Find Zav.

The children who heard this version about a doll were inclined to identify only that doll as “Zav”, and ignore other dolls. However they did not consider Zav to be the proper name of the block, perhaps already knowing that blocks are unlikely to have names. If they heard:

14) This is a Zav. Find the Zav.

then the children were inclined to treat Zav as a common noun, choosing several dolls or blocks as named by that word. This is one of the very few studies to test and to find children’s specific sensitivity to the meanings of morphemes that they did not yet produce reliably, not just to their presence as phonetic parts of the utterance.

Unfortunately, methodology is critical here. Standard tests (e.g. picture choice) of the comprehension of morphology have traditionally NOT suggested comprehension much in advance of productive mastery. For example, the classic work of Brown and colleagues as well as standardized test procedures typically suggest rather poor control of

the contrasting forms until well after age 3 or even 4 years (Fraser, Bellugi & Brown, 1963). Shipley, Smith and Gleitman (1969) lowered the task demands by asking 2 year old children to follow commands that varied in their grammaticality. In a later study with similar motivation, Gerken and McIntosh (1993) studied a group of 21- to 28-month-olds, varying whether the morpheme (e.g., the article) was supplied correctly:

15) Show the dog to me.

was missing:

16) Show dog to me.

replaced by another grammatical element:

17) Show was dog to me.

or replaced by a nonsense syllable:

18) Show gub dog to me.

Children had to choose which of four pictures was being talked about, and the other three pictures were not related in sound or meaning to the target. Nevertheless, even children with MLUs below 1.5 responded more appropriately with action to the complete grammatical utterances than to the other three options. Though even the least developed children were sensitive to misuses of the morphemes, this is still not evidence that they knew the meanings carried by the forms. Still, the expectation that the functional morphemes should be manifest seems to be present early on. Hohle and Weissenborn (1998, 2000) tested very young children's sensitivity to the presence of determiners in German sentences. Using a head turn preference procedure, infants as young as 7.5 months showed longer listening times to passages containing closed class items that they had previously listened to in a list, belying the assumption that closed class items were somehow phonetically unprivileged. Instead, it seems possible that children were forming lexical entries for closed class items to serve as possibly anchors to help identify open-class items such as nouns, a proposal that receives tentative support for infants over 10.5 months in Hohle and Weissenborn (2000) as well as in Shady (1996) and Shady, Jusczyk & Gerken (1998).

More effort is needed to demonstrate that the sensitivity to the functional elements extends beyond phonology, if the two research methodologies can be expected to converge. Shady et al (1998) found that 10.5 month-olds were sensitive to the phonetic

form of functors in English (ko kitten is hiding vs the kitten is hiding), but they were insensitive to their position until 16 mos of age (was kitten the hiding vs. the kitten was hiding). There is much new work in progress on this issue, at long last.

3.1.2 Parametric variation

Ironically, at the same time that evidence was mounting against the proposal of limited functional categories in child grammar, a move was afoot in Linguistic Theory towards minimizing unnecessary structures (Chomsky, 1992). It has been argued that there may indeed be parametric variation in the functional categories across languages (Fukui & Speas, 1986). At least three alternatives have been proposed:

Some functional categories may not be present at all in some languages.

Functional categories in some languages might lack a specifier position.

There may be variation in the feature set associated with functional categories.

Speas (1994) put forward a principle called Economy of Representation, that restricts the building of unnecessary structures hosting empty elements – a principle in keeping with the Minimalist position expounded in Chomsky (1995). Applying the principle to acquisition, children may obey Economy of Representation until they come across the necessary evidence that their grammars are insufficient without further structure (Roeper and Rohrbacher, 1994). And languages might vary in how early they supply that evidence. Suppose it is true that the grammar of Japanese lacks the C component. If so, then the “no CP stage” might be a kind of “default assumption” appearing in early grammars before children receive evidence that C is required for their language. Since German requires the verb to move to C even in simple declarative sentences, the evidence for CP in German is abundant, and small German children seem to use this verb position appropriately at a young age (Poeppel & Wexler, 1993; Verris & Weissenborn, 1992).

Radford (1995) modified his position to reflect the language difference, arguing that the morphological richness in Italian verbs, for example, obligates the early construction of a functional node, IP, to carry the inflectional morphology. As expected, this functional category appears earlier in the grammatical development of Italian children (Hyams, 1987; see arguments however about how perfectly inflections are

known in Pizzutto & Casselli, 1992; Hyams, 1992). Clearly not all children take the “no functional categories” route as a first step, but the Economy of Representation approach allows for cross-linguistic variation.

Deprez (1994) proposed that the delays in the area of functional categories reveal a position called the Weak Continuity Hypothesis, in which the functional structures are available to the child through UG, but are at first under-specified. As a result, derivations occur that are not compatible with the adult grammar to which the child is exposed, though again, not out of the range of UG possibilities. She further proposed that some parametric settings would be set early and obligatorily: Wh-movement and V-to-I movement are two that are argued to fall into that class. Since both involve operator variable relations, to leave them unmoved (either at S-structure or at LF, parametrically) would result in vacuous quantification, a serious violation of UG. However, V-to-I movement and V-to-C movement both involve specification of the strength of INFL, for example, for both N and V features. Hence delay may be necessary to avoid mis-setting a parameter.

On this view, echoed by several other theorists, acquisition of the functional categories consists in the acquisition of the feature contents of their heads, a process of lexical feature acquisition (Chomsky, 1995; Baauw, 2000; Hoekstra & Hyams; 1994; Grimshaw, 1994; Roeper, 1996). Significant ramifications in the rest of grammar result from underspecified feature content of functional categories, and the process is a necessarily slow one. But is an incomplete feature set compatible with UG and Continuity? Perhaps yes, if there is parametric variation across languages in these feature sets.

3.1.3 Maturation

If Radford's proposal (1990a) were correct, then children's first grammars reflect a "wild" grammatical option, one not represented in the world's languages and thus against Continuity. How could this be remedied? Wexler (1996) and Radford himself (1995) invoked maturation as the explanation. Radford's data were in fact all from children less than 2 1/2 years old. Perhaps then their immature sentences reflect a stage before true grammar emerges, a kind of proto-grammar. But it might then be expected

description of the debates about such a stage would take us well beyond the constraints of the present paper: it is mentioned here to show a further alternative in the functional categories debate, namely optionality. An excellent review is provided in Hyams (2000).

3.1.4 Summary

The no-functional-categories approach threatened Continuity, though it is not clear that any UG principles are violated by a minimal structure. One solution proposed is that the full structure is really there in the child's competence, if not in performance. At least, comprehension seems to be ahead of production. A second solution was to argue that reduced structures may be parametric alternatives and therefore default options in development. The notion of incomplete feature sets rescues Continuity by making analogy to lexical acquisition, which is necessarily input-dependent. The third alternative is optionality with complete structures, so far just argued for Tense, but potentially general. The question is: Do optional rules conform to UG?

3.2 Null subjects

One phenomenon that has attracted significant attention in early grammars cross-linguistically is the tendency of very young children to omit the subject of their sentences:

- 21) want go get it
 take a nap

English; Bloom, Lightbown and Hood, (1975)

- 22) veut lait (wants milk)
 est tomb´e (is fallen)

French; Pierce (1992)

- 23) blode mag nich (stupid (ball) like not)
 gleich macht fleige (soon will fly)

German; Verrips and Weissenborn, (1992)

- 24) voglio bere (want to drink)
 va a lavorare? (go to work?)

Italian; Guasti, (1994)

However, while this is ungrammatical in the adult grammar of English for sentences other than imperatives and certain styles such as writing in diaries (“Diary drop” Rizzi, 1994), it is perfectly legitimate in the adult languages that contain rich agreement morphology, such as Italian and Spanish. Hence the English child’s early grammar seems discontinuous with the adult grammar in this respect: young children allow null subjects where adults do not.

3.2.1 Performance considerations

The immediate possibility that comes to mind is that children are exhibiting a kind of length constraint on their early sentences, and constituents are dropped to reduce length (Brown, 1973). Such an account would contradict the notion that children’s productions are solely a product of their grammars, hence questioning Modularity. Several other findings must be mentioned in assessing the claim of a length constraint. It has been known since the first reports of this phenomenon that subjects are more vulnerable than objects (McNeill, 1970; Hyams & Wexler, 1993). Why would it not be the case that either one was dropped to conserve length? Secondly, the missing subject is retrievable from the immediate context, i.e. is referentially specific (Bloom, 1970). This suggests a semantic/pragmatic constraint in common with the languages that permit null subjects, not just a random omission. Notice though that this would be hard to prove: if the child said “Eat hay” out of the blue, would we be likely to recognize that he was attempting the generic proposition “horses eat hay?” Thirdly, children alternate production of sentences with null subjects, with sentences with fully specified subjects. Bloom, Lightbown and Hood (1975) gave examples of this in transcripts, and Braine (1973) referred to alternations as “replacement sequences”: clearly the length constraint is neither absolute nor insurmountable.

P. Bloom (1990) published a processing account of the null subject phenomenon

in which he proposed that for young English-speaking children, subjects are initially represented as full lexical NPs or as pronouns. Because of production limitations on young children's speech, that subject is subsequently dropped. Bloom claimed that a gradual change occurs over development in the likelihood of this subject being dropped. A corollary assumption was that lexical subjects are weightier and therefore more likely dropped from the output than pronouns. L. Bloom (1991) also invoked cognitive limitations to explain why children omit constituents, because subjects were more often missing with longer VPs. Valian (1991) put forward similar proposals, citing Gerken (1991) on the relative prosodic differences that cause differential vulnerability of subjects versus objects in very young children's imitation of sentences.

The performance accounts occasioned a rapid counter-response. Hyams and Wexler (1993) criticized the model on both formal and empirical grounds. Theoretically, it was unclear how dropping the subject might ease computational load, since it has to be represented before it was dropped. But more importantly, why is the subject more vulnerable than the object? Hyams and Wexler found the performance account unconvincing in that regard. In contrast, grammatical accounts can be provided for the asymmetry (see below). Empirically, the argument about lexical subjects did not stand up to the test by Hyams and Wexler: the proportion of lexical subjects remains roughly constant over time, but the proportion of pronominal subjects rises, as the null subjects are replaced. The one fact that remains in favor of the processing account is that the length of the VP is proportional to the heaviness of the subject, as if there were indeed a trade-off in the output (Bloom, 1990). However, the same property proved to be true of adult Italian (Hyams & Wexler, 1983), which is surely not due to output limitations!

The debate is not closed. Valian, Hoeffner and Aubrey, (1996) repeated Gerken's (1991) imitation study and showed that children below MLU 3.0 do exhibit performance limitations in imitating subjects. Pronominal subjects proved more vulnerable to deletion than lexical subjects, expletive pronouns more vulnerable than referential pronouns, and omission was more likely with long VPs. They argue for a full competence model of English children's subjects, and attribute all the variability to performance factors, with no acknowledgement of Modularity considerations.

3.2.2 Parametric variation

As an alternative to a performance limitation account to preserve Continuity at the expense of Modularity, consider the alternative in which the child's grammar is set to a different parameter from the adult's. Hyams (1987) proposed the first parametric account of the null subject phenomenon, based on a comparison of Italian and English with respect to the null subject parameter, that set off a flurry of other work. Hyams accounted for the facts above about the differences cross-linguistically, by using an analysis from Rizzi (1982) that in certain languages the agreement features contained in INFL constitute a particular kind of pronominal, namely *pro* (Chomsky, 1982), and this licenses ("properly governs") an empty subject in spec-INFL. This then constitutes a "pro-drop" language. However, in languages without such features, a pronominal subject must be used. Subsequent theoretical analyses retain the spirit of this parametric difference (Radford, Atkinson, Britain, Clahsen, & Spencer, 1999). In Hyams' account, English-speaking children begin with the parameter set to the Italian setting, namely permitting null subjects. In addition, auxiliaries appear in INFL in English, but are analyzed as main verbs in pro-drop languages. She argued that only when this parameter was reset, would auxiliaries be properly analyzed and begin to appear in the child's sentences. Thus one should not get modals in subjectless sentences.

The immediate problem then arises of learnability: how does the initial parameter setting get corrected? Hyams proposed an ingenious trigger, namely the existence of expletive sentences like

It is sunny

or

There is a new guy in my office

which reveal to the child that in English, subjects must be overt, even if they are semantically "empty". Hyams argued that once expletives are analyzed, then the parameter would get re-set to require overt subjects.

Unfortunately, it was soon pointed out that cross-linguistic data were incompatible with the idea of a universal trigger: in particular, Berman (1990) argued that

Hebrew contains expletives but also permits null subjects in some contexts but not others. She contended that Hebrew illustrates that the idea of parameters as all-or-none settings is too often an overstatement of the cross-linguistic facts (see also Vainikka & Levy, 1999).

Other properties have been proposed as triggers, for example Roeper and Weissenborn (1990) argued that hearing a null subject in an embedded tensed clause is a sure indicator that the language has the null-subject parameter setting. Jaeggli and Hyams (1988) argued that languages that allow null subjects are morphologically rich languages, whereas languages without it are morphologically impoverished: perhaps the child has to set a morphology parameter before the null subject parameter. On a more intuitive level, researchers have argued that the dangers of triggering are that parameters can be re-set on the basis of such limited data, say a single sentence. Of course that is what is supposed to be GOOD about parameters, but if the input data contain noise, that is a problem. For example, English, a non-null subject language, has plenty of sentences with missing subjects:

- 25) Got milk?
 Want coffee?
 See ya!

The danger is that the child would be in constant oscillation between parameter settings, the so-called pendulum effect (Randall, 1990; Valian, 1990).

Not just the theory but the data Hyams presented were also soon a matter of debate. Valian (1991) presented compelling evidence that English children's sentences at the early stages in fact do not look like Italian children's sentences without subjects, in a purely quantitative sense. Children below age 2.2 years in English dropped subjects about 31% of the time, whereas Italian children in the same age group drop 70% of their subjects, just as in adult Italian. Children in the null subject stage in English never produce null subjects in embedded sentences, but in Italian they do (Roeper & Weissenborn, 1990, Valian, 1991, Rizzi, 1994). Clearly the English child's early null subjects were a different phenomenon, restricted to the root clause.

Rizzi (1994) proposed that the null subject in children's grammar be analyzed as a null constant (Lasnik & Stowell, 1991). A null constant can only appear in the specifier of the root, that being the only position exempted from the identification requirement: it is

not c-commanded by anything, and it must achieve its identification via discourse. Via an elaborate chain of reasoning, Rizzi argued that in early Child English, the CP layer was truncated, allowing the top layer to be the IP. The SPEC of IP position was then an appropriate site for a null constant (Lasnik and Stowell, 1991), namely a root null subject. In this way, null subjects were argued to be discourse-licensed in early English. But at some cost: the price was the optionality of the CP layer. Rizzi accounted for early fronted wh-questions by arguing that they must be adjoined to IP, also proposed in de Villiers (1991). However, this creates yet another departure from adult grammar in need of explanation. Perhaps the higher node can be characterized as neither IP nor CP at first, but have features of each.

Both Rizzi (1994) and Hyams (1994) independently argued that null subjects should not be attested in wh-questions because the spec of CP is either occupied by an empty topic (Hyams) or missing (Rizzi). In support are data from Valian (1991) for English and Crisma (1992) for French, both of which reported very few null subjects in wh-questions. In contrast Radford (1994) and Roeper and Rohrbacher (1994) found many null subjects in Wh-questions. In the latter paper, this was linked to the presence of non-finite verbs. Specifically, Sano and Hyams (1994) and Roeper and Rohrbacher argued that before agreement is acquired in English, and therefore a functional head to house it, *pro* (i.e. the null subject) is licensed in the spec of VP, that being the most complete layer. In other words, these authors have a solution in which a functional category (AGR/Tense) is optional, as in Asian languages. But rather than a discourse linked null constant (Rizzi), they analyze null subjects as *pro*.

They argued that the null subject follows from a different parameterized property: there is not a *pro*-drop parameter *per se* in their model, but rather overt subjects are necessitated to fill the specifier position of AGR-SP in languages where that is not filled by an agreement affix. So, English being a weak agreement language, AGR-SP is at first absent, as it can be in Japanese. When the weak English agreement comes in, then AGR-SP is built to accommodate it by the assumption of Economy of Projection (Speas, 1994). From that time, the subject must be overt. Roeper and Rohrbacher argue that their model avoids the problem of negative evidence (none is needed) and also parameter-resetting—there is in fact no *pro*-drop parameter. Furthermore against the performance account of

Bloom (1990), they argue that the loss of pro-drop is only apparently gradual, rather it consists of a series of dramatic changes as the properties of agreement are mastered. So, for instance, in the child ADAM, immediately after the first cases of contrasting agreement, and again after agreement becomes productive, there is a big drop in null subjects with non-finite Wh-questions. On this model, then, null subjects are folded into the general theory of the economy of projection of functional categories. What remains unexplained on such accounts is the variability across sentences in the same transcript, and the departures of English from Italian children's proportionate use.

Hyams (1991) also proposed that English children might be adopting a Chinese type grammar in which subjects are optional on a discourse basis, not as the result of a grammatical parameter. In Chinese and Japanese, a null subject is connected to a discourse topic, rather than recovered from the rich verb agreement morphology as in Italian. However Wang, Lillo-Martin, Best and Levitt (1992) found Chinese children at the same stage also resemble Chinese adults in dropping subjects 55% of the time. In addition, Chinese allows the dropping of objects on a discourse basis, and Chinese children drop objects 20% of the time, like their parents, but English children only drop objects around 8% of the time. English children's grammars therefore resemble neither Chinese nor Italian in proportions of null arguments.

Lillo-Martin argues that American Sign Language adds to the picture of the possible parameter settings. Like Chinese, American Sign Language allows both subject and object to be omitted. But unlike Chinese, ASL is a rich agreement language in which the verb agreement contains the information necessary to retrieve the missing arguments, so it is not discourse optionality. She proposes that there are three possible parameter settings:

No null arguments – like English

Null subjects but not null objects – like Italian

Null arguments - like ASL.

Lillo-Martin's (1991) studies of ASL acquisition reveal some interesting stages: children acquiring ASL have many null arguments at first but without any verb agreement. They then acquire verb agreement for present referents before they acquire it for absent referents, presumably because of the deictic properties of the signs for agreement: it is

easier to point to a concrete entity than an abstract representation in signing space. It is only by age 5 or so that the full grammar of verb agreement and null arguments is in place. Does this mean that ASL learners move through a grammar like Chinese, or is the parameter right from the start, but the agreement morphology takes time to learn? A possible test arises from the following property of Chinese: when a topic phrase is fronted, it forms a barrier to the connection between the discourse topic and the null subject, so null subjects are not permitted in Chinese with, say, an object as the fronted topic (Yang, 2000):

- 26) *Sue e xihuan t
Sue (but not Mary) John likes

If ASL learners produce such sentences, this would be clear evidence that they do not have Chinese-type grammars, but instead have the Null Argument setting. Of course, then there is something else they do not know, namely the obligatoriness of the verb morphology. Discontinuity has to be posited somewhere else, but arguably in an area of the grammar where input matters and learning is inevitably protracted. In light of other cross-linguistic work that agreement morphology is virtually perfect from the start (Guasti, 1994; Clahsen & Penke, 1992; but see Pizzutto and Casselli, 1992) this remains a puzzle.

But do English children start with Chinese-type grammars? Yang (2000) makes the argument that some data are compatible with this possibility. Recall that Chinese forbids null subjects when an argument, the object, is fronted. In English this occurs with object wh-questions such as:

- 27) What did you see?

The prediction is that English-speaking children should never produce:

- 28) What see?

that is, they should always keep the subject in an object wh-question. Yang analyzed the first 20 files of the ADAM corpus in CHILDES and confirmed that prediction: 97.2% of the object wh-questions had the subject present. This provides important evidence that the null subject stage in English is not just performance-based, but adheres to

grammatical principles, albeit with properties at first of the wrong grammar. So perhaps Hyams was partly right: English children do get the parameter wrong at first, but it does not seem to be Italian that they choose. But neither is it decisively Chinese. As mentioned, for Valian, Hoeffner and Aubrey (1996), it is definitely not a wrong parameter setting at all, as their subjects' imitations matched none of the above competence models.

The argument that children's subjectless sentences are a product of their grammars, but that their grammars are set to a different parametric setting than the adults', can only work if the learnability problem is solved. Yang (2000) offers a radical alternative that combines the best of several models and avoids their problems. On Yang's account, the child has multiple grammars at first, all compatible with UG (also proposed in Roeper & Rohrbacher, 1994, Roeper, 1999; Kroch & Taylor, 1997). Yang proposes that the progress towards the adult model is essentially one of natural selection, that is, the fittest alternatives survive in a probabilistic manner. The model is in broad spirit like one proposed by Clark (1992), with several important differences. In Clark's model, the fittest grammar was chosen by computing how well all grammars parsed a large sample of sentences, with attendant huge computational cost. On Yang's model, each input sentence is analyzed by a grammar chosen with a probability, P , that then gets modified if the grammar is successful or not in providing an analysis of the sentence. Outputs are produced by the same probability-weighted grammars. The process is an entirely passive one, akin to that adopted in connectionist models, but the difference between this proposal and earlier proposals is that it is entire grammars that are in competition. The model of success and failure and weighted probabilities is an old tried and true one from pre-cognitive science psychology. Several aspects of the model seem attractive for theory

It does not require unambiguous evidence – a good thing since the existence of unambiguous evidence is questionable

it apparently does not suffer from the potential problem of parameter interference:

linked parameters rise and fall together if the whole grammar gets rewarded or punished, but converge on the target with time (Yang & Guttman, 1999)

the pendulum effect is real: the oscillation is among multiple grammars, not just two, but none of them are conclusively discounted by limited data.

the child's grammar at any one point consist of fragments of partial grammars, not

necessarily just one from the set of UG grammars.

Notice that UG is built into the assumptions of the model, so this is not a return to Gold's theorem (1967) in which the original state is ignorance of language, nor is it a theory-neutral connectionist-style model. Much work remains to see whether it is realistic in its time course predictions: will the multiple grammars converge in sufficient time? It maintains the possibility that in many adults, multiple grammars might still reside (Roeper, 1999): a considerable comfort to those of us exasperated by adult linguistic variability. Importantly, the Multiple Grammars approach has a different account of how optionality of elements is replaced by obligatoriness, which on Wexler's (1999) account must necessarily be by maturation, as it is not learnable. On the Multiple Grammars approach, the two or more grammars in competition provide the variability: as the competing grammars become lower in probability, the forms become obligatory. Note that the unsolved problem of variable inflectional acquisition is also accommodated under this idea without invoking separate learning mechanisms. However, the idea that at any one time the child's grammar consists of fragments of many partial grammars would seem on the surface at least to contradict other proposals about systematicity in early grammars, such as Weissenborn's (1994) Local Well-formedness Constraint hypothesis, to the extent that it applies across utterances.

It will be interesting to compare the reception in the field of the multiple grammars model with the reception to Valian (1994)'s related proposal about the child weighing the success of parameter settings. However, there are several differences to be noted. Valian (1994) proposed that the child might maintain both settings of the null subject parameter “*on an equal footing* until sufficient evidence accrues to favor one over the other” (italics mine). She argued that the child's parser, driven by her grammar, would be unable to handle a sentence containing the opposite parameter setting. As a result, if the child set the parameter prematurely, she would never analyze the evidence needed to shift the parameter to the other setting. Hyams (1994) criticized this assumption by arguing that it was precisely failed parses that drove the parameter setting model (Clark & Roberts, 1993):

“ a sentence a is a trigger for a parameter value P_x just in case a grammar must have P set to x in order to assign a well-formed representation to s ”

Valian proposed instead a scale model in which all of the parameters were kept in limbo until enough evidence had been accumulated. Hyams (1994) dismissed this as follows:

“in order for the child to parse all the triggering data for all parameters, he must start out with the entire set of possible adult grammars – a rather implausible assumption, on the face of it” (p. 296).

If the child has the union of even two grammars, what will allow him to reject this “super grammar”? Hyams argued that the computational resources were also too great to be realistic: the child would have to remember previous input data, past failed hypotheses, as well as the ability to represent and compare competing representations of a sentence generated by competing grammars. Furthermore, the younger the child, the larger the number of grammars that would have to be held in memory.

These criticisms may not apply to Yang’s model. The child does not maintain all grammars “on an equal footing” until evidence accrues to decide between them: the process is a dynamic one, and their weight and probabilities of selection continually shift. Furthermore, the model is a passive one, without requiring unrealistically sophisticated memory and representational resources. If connectionism research has done the generative paradigm a favor at all, it is making clear that networks, armies of enough dumb units, can achieve extraordinary success without higher level computational resources. So the multiple grammars model is not the scale model in important respects.

The Multiple Grammars approach represents a retreat, of sorts, to a model highly dependent on input, and involving a potentially protracted period of uncertainty in grammars. But it retains Continuity: any given utterance must be compatible with UG, and the principle of Modularity: syntax is modular and insulated. Those properties may allow the paradigm to accommodate it.

Notice, however, that the Multiple Grammars approach threatens the Modularity Matching model as a methodological imperative. Under the view that at any given time the child’s behavior might be a function of multiple UG-constrained grammars, it is unlikely that performance in even the best contrived experimental setting will conform to

the expectations based on any one grammar. This has the unfortunate effect that it may allow even more degrees of freedom than the competing factors models allowed. Perfection may be beyond reach.

3.2.3 Maturation

No specifically maturational account of null subjects has been proposed. In other words, the above accounts assume changes in functional structure or parameters are due to evidence, not to time-related growth. Furthermore, the cross-linguistic data would seem to contradict a simple maturational approach. Children acquiring different languages show considerable variability at the earliest stages in their proportions of subjects (Valian, 1991). The data do not seem compatible with a model in which certain principles or structures only come on board after a time lag, or in which optionality is resolved by time alone.

3.2.4. Summary

There seems to be developing consensus that child null subjects are a grammatical phenomenon. The existing accounts disagree as to whether there is a specific parameter for null subjects, whether null subjects are parasitic on some other parameter to do with features in INFL, concerned with weak versus strong Agreement, or whether children's null subjects should achieve a different interpretation, say under discourse agreement as in Chinese. The facts seem compatible with a Multiple Grammars approach, but how widely that will be accepted as a model remains to be seen.

3.3 Principle B

A huge amount of attention has converged on a well-documented failure of young children to observe Principle B of the Binding conditions (Chomsky, 1981) , namely:

If an NP c-commands a pronoun within the same clause, they cannot be co-indexed.

This principle is what rules out the interpretation of:

29) Bob likes him

as meaning Bob likes Bob, himself. In many experiments across several languages, however, the consistent finding has been that children allow that reading until surprisingly late, even at six years. The basic finding can be demonstrated in an extensive study by Chien and Wexler (1990). Children aged 2;6 to 7 were given yes/no questions about a series of pictures that had various actions depicted between two characters, say Papa Bear and Baby Bear. The proportion of children answering YES to the question:

30) Is Papa Bear touching him?

when Papa Bear was touching himself, was 70% at age 3, and still 24% at age 6-7. The responses to reflexive pronouns such as:

31) Is Papa Bear touching himself?

were in contrast, uniformly good, as were correct YES responses when Papa Bear was touching Baby Bear and the pronoun was used appropriately. A simple confusion of the pronouns and anaphors won't work as an explanation, because children almost never confuse reflexives for pronouns.

The immediate conclusion might be that children do not know Principle B until later, but then significant questions about learnability arise. However, some additional results from Chien and Wexler (1990) suggest children do obey the Principle in other kinds of sentences. For example, when the sentence contains a quantifier:

32) Every bear touched him.

children who fail on (30) nevertheless reject (32) as meaning every bear touched himself. Chien and Wexler interpreted this result to mean that children DO know Principle B, but do not obey it in certain limited circumstances, namely with a referential NP as in (30). Thornton (1991) noticed a similar phenomenon using the Truth Value Judgment task instead of yes/no questions. She set up a story with different characters in which different things are acted out, for example, Big Bird covering Grover and Papa Bear covering himself. Then a puppet is made to say:

33) I know who covered him. Papa Bear.

Again, children who showed the perplexing failure on referential NP's, successfully

rejected the truth of the sentence with the embedded question. So children cannot be said to lack Principle B entirely.

3.3.1 Performance considerations

Grimshaw and Rosen (1990a,b) offered a pragmatic interpretation of this strange array of facts. They contended that intonation and context may play significant roles in the assignment of readings to pronouns in sentences such as (30). For example, it is possible to stress the pronoun and get an inkling of a co-referential reading:

34) Papa bear covered HIM

yet such a reading is not available for sentences such as:

35) Every bear covered HIM

(see also McDaniel & Maxfield, 1992). Grimshaw and Rosen argued that in the presentation of the task there is no natural antecedent supplied for the pronoun, so the child is forced to choose between violating Principle B or violating the discourse requirement on antecedents for pronouns. In fact even young children very rarely choose unmentioned objects in the context as possible pronoun referents (Wexler & Chien, 1986). So Grimshaw and Rosen concluded that the children's knowledge is masked by various interfering factors. What is slightly odd about the claim is that in fact many studies that have found the result DID mention the participant in the just preceding sentence e.g.

36) Here are Papa Bear and Baby Bear. Is Papa Bear washing him?

Grimshaw and Rosen tried to explain the discrepancy between quantified sentences and non-quantified sentences in performance terms. They proposed that the bound variable interpretation between operators and pronouns may not be available to children for (32), leading them to focus on the only other possible reading, that of the deictic reading, hence getting it right by accident. That is, it is the quantifier sentence that is unambiguous, whereas the referential NP sentence is ambiguous because of potential intonation and contextual biases. The suggestion about the lack of bound variable

readings is belied by the considerable evidence of such interpretations available for both quantifiers and wh-questions by very young children (de Villiers & Roeper, 1993). The Grimshaw and Rosen account seemed to require too many *ad hoc* assumptions, and to posit a different kind of knowledge failure to account for the contrasts across sentence types that does not ring true.

Grodzinsky and Reinhart (1993) made a different suggestion, founded on the alternative theoretical conception of Principle B in Reinhart (1983, Reinhart & Reuland, 1993). Reinhart had proposed that Principle B only applies to pronouns bound by an operator, hence does not apply to pronouns co-indexed with a referential NP. On this theory, co-reference is ruled out by Principle B in sentences such as:

37) Who touched him?

or

38) Every bear touched him.

but not by Principle B in:

39) Papa Bear touched him.

This is an attractive account for the child data, because children's data discussed above then reveal precise knowledge of Principle B! But what then, rules out our adult interpretation of (39) as co-referential? On Reinhart's scheme a pragmatic rule, Rule I, governs how speakers and hearers figure out co-reference in situations where the syntax allows more than one possibility. Rule I is a form of economy, that states that establishing a referential dependency in the syntax, e.g. by binding, is more economical than doing it in the pragmatics, via co-reference. In the particular case of (39), the hearer must compare the two alternative construals of the sentence, one derived via binding and one via co-reference. This amounts to deciding whether a reflexive would be allowed in the position of the pronoun. If it would be, but was not chosen, then the reading of the pronoun must be intended to be deictic. This could be considered a variety of a Gricean maxim, namely, to "avoid ambiguity" (Crain & Thornton, 1998), but economy principles are also involved. It might seem reasonable to conclude that children know Principle B, but do not yet know Rule I. However, that is not the move made by Grodzinsky and Reinhart. They claim that the child knows Rule I as well, but lacks the processing ability

to carry it out, because it involves maintaining in memory both the original representation of the sentence, and the alternative, with a reflexive, in order to carry out the comparison. This is a subtle account terminating in a performance account of the Principle B failure, that preserves continuity of grammar (and even pragmatics) at the expense of positing a processing limitation. Children are said to know binding, but have problems in their performance with co-reference.

The delay in observance of Principle B is not, however, universal across children learning all languages. How could this be covered under a processing limitation account? The difference across languages seems to relate to whether the pronoun can be a clitic or not: languages in which weak pronouns appear as clitics seem not to show the Principle B delay (Philip and Coopmans, 1996; Baauw, 2000). McKee (1992) attributes the exemption of clitics to their movement properties, e.g. that the clitics move outside of the VP to attach to INFL:

40) Italian: Papa lo lava
 [Papa him washes]

McKee claims that English children might misconstrue the VP as the minimal governing category for the object pronoun, and hence allow it to corefer with the subject (outside that domain). Because in Italian the clitic moves up to INFL, there is no possible misconstrual of what the minimal governing category is, and since the clitic and subject are both in it, coreference is ruled out by Principle B. Unfortunately, McKee's account fails to handle the operator binding facts above, and was also argued to be incompatible with the notion that the subject begins internal to the VP (Avrutin & Wexler, 1992).

Avrutin (1994) attributes the difference between clitics and pronouns to their referential properties. Clitics cannot be used to refer deictically, i.e. in accompaniment with a pointing gesture. In English, one can say:

41) Bill likes him

and accompany it with a point at "him", but in Italian one cannot point at the object with a clitic pronoun, just with a strong pronoun. Avrutin argues that only pronouns that refer deictically can co-refer, ruling out clitics as potentially co-referential.

3.3.2 Parametric variation

Early in the research on binding, the possibility was explored of parameterized options for Principle B. Jacobowicz (1984) had proposed that perhaps the Subset Principle could be used to predict that children would make the mistake of treating pronouns as anaphors, thus potentially accounting for the above mistake. Wexler & Manzini (1987) discussed in detail the setting of parameters for binding theory. The Subset Principle set the problem of parameter setting and the lack of negative evidence in the following way: if one language is a strict subset of another language, i.e. contained in it, than the child's optimum learning strategy would be to start with the parameter set to the smaller language. Then positive evidence would arise that contradicted that setting of the parameter, and it could be reset. In this way, the Subset principle could specify a markedness hierarchy for parameter values. So, for instance, locally bound anaphors (as in English) were argued to be unmarked relative to non-locally bound anaphors, such as are found in Icelandic. But Wexler & Manzini also demonstrated that locally bound anaphors are not unmarked with respect to pronouns because the Subset Condition is not met: one type of language is not a subset of the other. Hence Jacobowicz (1984)'s results could not be explained by the Subset principle.

An attractive proposal about cross-linguistic variation is made in Baauw (2000). He argues that there are two types of clitics, phonological and syntactic, and uses Dutch to try to distinguish which properties of clitics make them relevant for the exemption from Principle B delay noted above. Dutch weak pronouns allow a separation of the various properties, and Baauw's results suggest that Dutch weak pronouns, even though they cannot be used deictically, do show the familiar delay. Avrutin's proposal about deixis is therefore too strong. It is critically movement of the clitic that creates the exemption from the delay in Principle B obedience, that is, children do well on Principle B with moved clitics. Why is movement relevant? Recall that McKee proposed that it permitted children a proper determination of the governing domain. Baauw provides a more elaborate account in which he argues that moved, syntactic, clitics must be bound variables, and hence covered under the Grodzinsky and Reinhart (1993) proposal. But Baauw also emphasizes a further issue of cross-linguistic variation: Dutch children fail to recognize the weak pronouns as clitics because their status is ambiguous in the input. In

contrast there is ample evidence in Spanish that the clitics have moved, and Spanish children obey Principle B from the beginning. The rich morphology of Spanish and certain other languages (e.g. Norwegian) makes the status of the elements as syntactic clitics unambiguous, whereas in others (Dutch, Icelandic) there is uncertainty and hence a delay of Principle B. Significantly, we see that all the exceptions to the delay of Principle B are accommodated under one scheme, but the basic phenomenon remains: children allow too much co-reference between referential NPs and non-clitic pronouns.

3.3.3 Maturation

To an outsider to the paradigm, the amount of time and effort spent to deny that this result means that children do not know Principle B must seem peculiar. But Principle B is a fundamental principle of Universal Grammar, and the idea of Continuity is paramount. The alternative, namely that the child discovers Principle B on the basis of positive evidence, has generally been considered impossible. If Principle B just includes the relations between operators and bound variables, however, the data are compatible with early knowledge of it. Then the overly inclusive co-reference has some other cause, and a number of theories point to aspects of still inadequate pragmatics or processing limitations, with no clear consensus. It is interesting, though, that both Continuity and Modularity have been preserved as the story changed. Continuity of the grammatical principle remains intact, though at the expense of discontinuity or at least delay somewhere else. Modularity is preserved by extracting co-reference from binding: the influence of pragmatics and/or processing is not on the principle B but on some other component, i.e. however it is that listeners figure out co-reference. The field has learned a lot in the process, in that researchers agree what it is NOT, but there is still the puzzle ten years later for why children allow co-reference for:

42) Papa Bear washed him.

which is where it all started. Avrutin (2000) attributes the failure to a processing capacity difference between children and adults. To succeed in understanding pronouns, the child must coordinate his complete syntactic knowledge of Principle B with his developing

knowledge of discourse constraints. Avrutin emphasizes that knowledge of conversational rules requires the ability to go beyond your own knowledge and take into account the representations of other speakers in the conversation. He proposes that children (and aphasic patients) fail to make appropriate inferences about other speakers' representations of the discourse, and allow deixis with definite NPs without pointing. Avrutin claims this reflects a processing capacity difference from normal adults. This latter option is not open to Wexler if he is to maintain the strongest version of Continuity in which processing resources are the same.

Wexler (1999) proposes a maturational lag in the child's knowledge of the interface condition, namely the integration and co-ordination of syntactic knowledge—which is complete, for Principle B, and the conceptual discourse system, which is less adequate at first. Wexler (1999) has consistently argued that the child knows Principle B, as shown by the behavior with bound pronouns. Furthermore, the child knows what Wexler calls the “grounding principle”, namely that an NP must be grounded, either by a proper name, or deixis (pointing) or by being coindexed with another grounded NP. This latter is purely syntactic, but deixis involves subtle coordination of the attention of hearer and speaker, outside the domain of grammar *per se*. Wexler argues that the child might mistakenly assume that an NP is successfully grounded when it is not. What the child lacks is the “ability to make inferences about what speakers and listeners can infer in discourse situations” (p. 88). This latter is presumably not part of the pragmatic module of grammar, else it would be covered by Continuity and Modularity. Hence it is available as a last resort explanation for the co-reference failure.

3.3.4 Summary

The Binding principles constitute a clear example of UG-constrained knowledge: if Continuity is true, then the child should show early knowledge of these principles. At first it seemed as if children violated Principle B until quite late: now it seems agreed that the failure lies elsewhere than the grammatical principle. But there is an attendant cost: children's comprehension is held to be a product of the grammar in interaction with something else, and is Modularity not threatened? It depends on what the “extra” something is that the child has to acquire: processing capacity (awkward for both

Modularity and the strong continuity idea), Rule I (maybe Pragmatics is learned?) or the interface conditions (where Modularity, apparently by definition, does not apply).

3.4 Barriers to movement

A steady stream of research has pursued children's understanding of the constraints on movement rules first analyzed by Ross (1967) and subsequently central to linguistic theory in the generative tradition. In Chomsky (1985), the constraints were re-conceptualized as barriers to movement, and much subsequent work refined them in terms of licensing conditions on empty categories (Rizzi, 1990; Cinque, 1990; Szabolsci & Zwarts, 1992). They are among the most highly respected central tenets of the generative tradition, and used as examples of the poverty of the stimulus for learning, and the abstractness of the knowledge that children must eventually attain. They are often used as illustrations of the irreducibility of syntactic phenomena to e.g. parsing constraints or discourse, though this is not a closed question (Van Valin, 1998).

Phinney (1981) and Otsu (1981) were the first to ask whether children respect the barriers to a wh-question, using an innovative methodology. Everyday speech is impoverished in long distance wh-questions, so although no evidence had been adduced of young children disobeying the constraints on movement, neither was there much evidence that they allowed long distance movement at all. Furthermore, it was not feasible to present children with ungrammatical sentences and see whether they could get a meaning from them: not only might there be ethical considerations against such research, it would probably show that children can get meaning by stepping outside the grammar if forced to do so. Otsu (1981) tested whether children obey constraints on wh-extraction from relative clauses and prepositional phrases in a comprehension task. The innovative procedure used was to present very short stories to the children that set up two possible readings for a question, only one of which had legitimacy as a grammatical option. He presented his preschool subjects with short pictured stories such as:

43) Jane is drawing a monkey with a crayon. The monkey is drinking milk with a straw.

followed by a wh-question that was ambiguous only if the child lacked the constraint on extraction from a relative clause:

44) What is Jane drawing a monkey that is drinking milk with?

Adult speakers of English must construe the question with the main clause. If the child showed no constraints on extraction, the child might answer "a straw" on some occasions. Otsu's results were interpreted to show that children respected the barrierhood of the relative clause or the prepositional phrase. Nevertheless, the results were only statistically significant: there were rather more violations of the barrier than might be expected from obedience to a principle. However, he also included a test of comprehension of the relative clause, and found that the children who showed mastery of the structure in comprehension were more likely to respect the barrierhood of the relative clause. In other words, maybe one could only expect the barrier principle to apply if the children successfully analyzed the structure of the relative clause.

However, other possibilities arise about the children's responses in Otsu's task. In subsequent work, a systematic distinction was made between argument wh-words and adjunct wh-words. Chomsky (1985), Rizzi (1990) and Cinque (1990) discuss this in different terms, but in general the object argument wh-trace is licensed by lexical government from the verb, which means it can be co-indexed with the moved wh-word even at a distance and even across the barrier of an adjunct wh-medial:

45) Who_i did you ask how to contact t_i?

I asked how to contact your plumber.

The adjunct wh-traces in contrast must be licensed by antecedent government, and that can be blocked by an intervening wh-word such as the medial:

46) When_i did you ask how to make contact *t_i?

*I asked how to make contact next summer.

Subsequent research with 4 to 6 year olds (de Villiers, Roeper & Vainikka, 1990; Maxfield & Plunkett, 1991) revealed that children are sensitive to the argument/adjunct distinction and to these differential possibilities of movement of each type of question, in keeping with the licensing differences.

Otsu used exclusively argument questions (actually arguments of prepositions) in his research and the results suggest that children may be more lenient than adults in licensing the movement of argument questions. De Villiers and Roeper (1995b) asked whether children in the same age range would show better obedience to the relative clause boundary if adjunct questions were used instead. Adjunct questions cannot be extracted from within a relative clause:

47) How did the man who hurt his leg *t get home ?

In their experiment, they studied a variety of types of relative clauses: object and subject relatives, and even extraposed relatives (motivated and pronounced appropriately!) such as:

48) How did the dog climb the tree who barked?

In two studies, one cross-sectional and the other longitudinal, children from 3 years and up respected the barrier of the relative clause across all conditions and within a tolerable level of statistical “noise” (10% at most error).

A distinction between complements and adjuncts of different sorts has also been found to be systematic for young children (Roeper & de Villiers, 1992; Goodluck, Foley, & Sedivy, 1992). That is, young children do not allow wh-extraction from in-order to adjuncts, nor from temporal adjuncts, just like adults. However, in this case, the technique was a simple preference procedure, potentially subject to the criticism ii) in section 3.4.1 below.

3.4.1 Performance considerations

Otsu’s (1981) results are controversial in retrospect on two grounds:

Hamburger and Crain (1982) and Crain and Thornton (1998) believe that failures on relative clause comprehension in young children are artifacts of experimental procedures, so young children of 3 and 4 should be in perfect control of the

structures.

Crain and Thornton (1998) object to studies that show only statistical differences between grammatical and ungrammatical conditions. In particular, they argue against the preference method in that it is inconclusive about what the child allows or disallows. Biases in the context or story might lead the child to give one answer rather than another, and this might masquerade as obedience to some principle.

The objection to using preference as an index of adherence to grammatical principles is addressed in other experiments on barriers to wh-movement, that is, whenever it is possible to engineer a minimal contrast. Beginning in work published in 1990 (de Villiers, Roeper and Vainikka), a stream of research on movement barriers has used a common methodology in which an ambiguous story sets up possible construals for questions, and then is followed by a question. Sometimes the question is genuinely ambiguous: no grammatical principle restricts either answer. By observing children's preferences in that condition, we can tell whether the story or the situation is introducing non-grammatical bias. So far, this mirrors Otsu's work. However, a second version of the question is given to a second group of children (usually, a within-subjects design with counterbalancing of sets), and in this case it contains a barrier, say a medial wh-word. Nothing has changed about the story or the situation, so any change in the pattern of answers, say strong avoidance of the ungrammatical reading, is evidence of a barrier effect. In de Villiers and Roeper (1996), we call this the method of supplying context but removing it as a variable. That is, we try to satisfy felicity conditions for the question, but rule out the possibility that the conditions themselves determine the choice of answer by keeping them constant as the question changes in a minimal way. For example, take the following:

49) **Story:**

This little girl went shopping one afternoon, but she was late getting home. She decided to take a short way home across a wire fence, but she ripped her dress. That night when she was in bed, she told her mom, "I ripped my dress this afternoon on the fence".

Question:

50) When did she say she ripped her dress?

There are two possible interpretations of the question, depending on where the listener interprets the trace to be for the wh-question *when*- is it connected as an adjunct to *say*, or to *rip*?:

a) When_i did she say **t_i** she ripped her dress?

OR b) When_i did she say she ripped her dress **t_i**?

That is, the answer could be: *at night* (if a), or *that afternoon* (if b).

But the question:

51) When did she say **how** she ripped her dress?

allows only one reading, *at night*, because the medial wh-word blocks the trace in the lower clause.

de Villiers, Roeper and Vainikka (1990) found strong evidence that:
long distance movement is available to 3 year olds
the argument/adjunct distinction is respected
medial wh-words block long distance movement of questions for children just as for adults.

In the research on strong wh-islands (de Villiers et al, 1990) a striking finding that could not be easily accommodated was that young children commonly answered the medial complementizer as if it were a question. So, in response to

52) How did the boy say who fell?

they would answer the question “who fell?” e.g. “his sister”.

Could this be a performance error, that is, might the children just be answering the last question word they hear? Several control experiments rule out various non-grammatical explanations, for example, children never answer the complementizer “who”

in a relative clause sentence (de Villiers & Roeper, 1995b). Furthermore, children at the same stage produce the forms in elicited production tasks (Thornton, 1991), and judge them to be grammatical in judgement tasks (McDaniel, Chiu & Maxfield, 1995), so the "mistake" seems to be competence-based.

3.4.2 Parametric variation

In accounting for the medial error above, de Villiers et al (1990), Roeper and de Villiers (1994) proposed that children might be using the kind of grammar found in some dialects of German and other languages (McDaniel, 1989) that permits the question in a medial position scope-marked by an initial wh-question of usually invariant form. In German, it is possible to have a question such as:

- 53) Was hat er gesagt wie er den Kuchen machen kann?
 What did he say how he the cake make can?
 [How did he say he could make the cake]

Notice that the first question word "was" [=what in English] is not answered; it just serves to mark the sentence as a question. It is only the medial question "wie" [=how] that needs an answer.

Critical to our argument is the fact that partial movement occurs in the languages that allow it only where there is no subcategorization. In those languages, the medial question is the real question, not a complementizer, and the front question word is just a marker that the utterance is a question, but contributes no meaning.

However, we attribute the choice of this grammatical option to an immaturity in the child's grammar of subcategorization, perhaps in features governing c-selection. C-selection is lexically linked for complements, in that relatively few verb types take complements, and within that class there is considerable lexical variability, for example with factive verbs. Roeper and de Villiers (1994) argued that children at this stage do not appropriately subcategorize the lower clause. We argued that the child's complement

clause is attached at a higher level than for the adult: it is not syntactically subcategorized by the verb though it is thematically governed by it. So children's grammars permit partial movement structures primarily because they do not yet have the embedded question subcategorized under the main verb, though they recognize that it is thematically connected to it. This means that the verb in question requires the satisfaction of a particular thematic role such as theme or goal (e.g. that *say* does not stand alone as an intransitive verb). Notice that we do not have the option to propose that the embedded question be a true adjunct, or long distance movement would not be possible out of it. Clearly, a child may make a basic differentiation between adjuncts and complements and yet still lack the full syntactic subcategorization of a mature grammar. In subsequent work, (de Villiers, 1999), I have argued that the child's complements might lack a crucial feature relative to adult complementation, namely, a feature on the CP that permits the embedding of a proposition that is not true: a false complement. Children below the age of about 3 and 1/2 do not seem able to represent the meaning of a sentence such as:

54) He said/thought he saw a unicorn

at least when the child knows that there was no unicorn in sight. This point is returned to below, in the implications for modularity. The coincidence of the medial answer and the prevalent comprehension error with false complements seem to reflect difficulties of complementation and are roughly related in time, but they do not resolve together: the medial answer is retained longer (de Villiers & Pyers, submitted; Abdulkarim, 2000).

These analyses allocate the errors children make to incomplete feature specification in the CP, and suggest a lengthy period of lexical learning might be needed to correct it. Alternative analyses of the medial question phenomena are provided by Thornton (1991) and by McDaniel, Chiu and Maxfield (1995). Thornton put forward the claim that the child has full knowledge of complementation except for a parametric setting involving agreement between the spec of CP and its head. Thornton (1991) discovered that children's productions can also contain a "copy" of the initial wh-word when they are put in an elicited production situation and encouraged to produce complex sentences. For example, they produce questions such as:

55) What do you think what's in her hat?

and less frequently, cases in which the medial question is not a direct copy:

56) What do you think where the marble is?

Thornton's analysis also connected the phenomenon to the partial movement possibility in German, but she argued that the phenomenon is due to the child's rules for specifier-head agreement. That is, for a case of long distance movement when there is a trace in the intermediate specifier of CP, the children's grammar requires an explicit complementizer to carry agreement features, and that complementizer is realized as a *wh*-form. In support of this claim, children in her study did not produce these medial forms with infinitival complements, only tensed complements, because only the latter trigger spec-head agreement. In their later development, children conclude that spec-head agreement is only necessary for subject extraction questions, and that the agreeing complementizer must in fact be null (Rizzi, 1990).

The analysis does not cover the whole range of facts (see McDaniel et al, 1995) nor the fact that children answer the medial question for both infinitival and tensed complements (de Villiers et al, 1990). McDaniel et al, using primarily judgment data from young children, were able to trace the development of grammars permitting partial movement, and present convincing evidence that it is a stage through which children pass as they acquire full competence in English. They attribute it to the lack of a feature in the spec-CP, arguing that the complement is then structurally on a par with a relative clause. This does not mean that the clause will be interpreted as a relative clause: it refers only to a restriction on which elements might occur in spec-CP and CP. The feature to be acquired is one Rizzi (1990) called [pred] for predicate, and acquiring that feature allows children to make a structural distinction between *wh*-complements and relative clauses.

There is thus convergence of opinion that partial movement is represented in young children's grammars as an option. McDaniel, Chiu & Maxfield (1995) proposed a particular deductive trigger that could result in the establishment of the [pred] feature in children's grammar. Roeper and de Villiers (1994) argued that since moving away from partial movement depends on getting the particular subcategorizations right for particular verbs, the process could be prolonged and lexically specific. It is not yet clear how the Multiple Grammars approach (Yang, 2000) can be helpful in explaining the late occurring errors that may be due to other parametric options in complex grammar. If, for example,

the grammar of German is still being chosen to analyze or produce medial questions, with a certain fairly high probability, then why do we not see more evidence of other German structures in the English of three and four year olds? Surely the grammar is not just selected to deal with medial questions? Yang (2000) argues that it is implausible that the learner makes choices as to the relevance of particular parameters in analyzing an input sentence. Yet an outcome of his approach is that children will pass through stages in which some parameters have been fixed before others. Roeper (1999) discusses these questions in his paper on “Universal Bilingualism”, in which he invites a broader look at the existence of “radically different islands of grammar variation” even in adult language, that may reflect grammars never completely eradicated, but perhaps linked to lexical or register variation. He proposes that some grammars may never be completely removed because they represent defaults: “Two grammars will not assimilate if it requires the elimination of a more economical representation in either grammar.” Hence the medial answer may represent such a default option for English.

Abdulkarim (2000) undertook a comprehensive analysis of the medial question phenomenon to decide among competing alternatives that arose since the earlier syntactic treatments in McDaniel (1989). So, for example, Fanselow and Mahajan (2000) argue that the medial *wh* occurs as a default option for long distance movement, and as a consequence, should itself be subject to the same barriers that long distance movement is subject to. For example, negatives should block both long distance movement and medial *wh*-movement in a sentence such as:

57) Why did the boy not tell his mom how he broke the bike?

These stories, as you might imagine, are rather hard to write, given all the potential answers that have to be accommodated, but Abdulkarim’s results with 3 to 6 year olds bear out the prediction beautifully. In a carefully matched design, so that equivalent stories could be compared, children were prone to answer the medial question for a question like (58) but not with the negative (57):

58) Why did the boy tell his Mom how he rode the bike?

The basic results concerning strong islands have been replicated across children speaking several languages (see Roeper & de Villiers, 1994 for a review) using the same

methodology, and the methodology defended and described further in de Villiers and Roeper (1996). It remains to be seen whether the results can stand as a genuine universal, or whether there might not be parametric variations in this domain too.

A puzzling set of results that remain outside the current analyses concern children's sensitivity to other "weak" barriers such as quantificational adverbs (Philip & de Villiers, 1992) and negation (Abdulkarim, 2000). This work also used the methodology of contrasting questions following identical stories, and again, barrierhood is generally respected. However, in the case of quantificational adverbs, Philip and de Villiers found that children only respected the barrier of the adverbs once they knew their semantic properties. That is, quantificational adverbs such as *always* and *often* introduce special entailment characteristics: the sentence

59) John eats grapes with a toothpick

entails the superset:

60) John eats food with a toothpick

but with the adverb:

61) John often eats grapes with a toothpick

this does not imply

62) John often eats food with a toothpick.

It was only when children knew these semantic properties of the adverbs that they then blocked wh-questions from clauses containing them:

63) Why did John often say he liked ballet t*?

It seemed promising that this might provide a way to investigate the linguistic question of whether the weak island effects of adverbs, negation and factivity, were semantic or syntactic (Szabolsci and Zwarts, 1992). Negation subsequently failed to lend itself to such an analysis (Philip and de Villiers, 1993), and appears to behave as conclusively as a syntactic barrier (Abdulkarim & Roeper, 1997; Abdulkarim, 2000). Factive verbs, on the other hand, do not act as barriers for young children's wh-movement (Roeper & de Villiers, 1992; de Villiers, Curran, Philip & DeMunn, 1998). However, acquisition of their semantic properties (presupposition, monotonicity) seems to be on a path curiously independent of the acquisition of their barrier properties (de Villiers, et al 1998), and acquired surprisingly late in childhood (7 or 8 years). Much work remains to make sense

of the weak island phenomena and the extent to which they involve semantic information.

An additional contingent development of barrierhood appears in testing wh-movement with light verb constructions (de Villiers & Roeper, 1995a). In this study, children permitted extraction from both the following:

- 64) How did they make the decision to run t?
- 65) How did they like the decision to run *t?

“make the decision” is a light verb construction, and de Villiers and Roeper make the case that the determiner is in an NP rather than a DP position. They connect that to the observation that co-reference is not possible in the first kind of construction for adults:

- 66) He made the decision to shave him

but it is in the second:

- 67) He liked the decision to shave him

presumably because the DP introduces a new binding domain for the pronoun.

Interestingly, only those children who treated the two sentence types differently with respect to co-reference also treated them differently with respect to movement barriers. However, the results were considerably weaker than in the other barrier experiments, despite balanced story designs.

Baauw (2000) repeated the study in Spanish, which also has the distinction between light and strong verbs. However, Baauw argued that the appropriate distinction was not between analyzing the two structures as a NP versus a DP, but rather that both were DPs, with the weak one lacking certain features. He argues that children may initially misconstrue determiners as expletive determiners, a class that appears in several European languages. Zubizarreta and Vergnaud (1992) and Longobardi (1994) argue that expletive determiners occupy DP, but lack denotational content, being just bundles of phi features. According to Baauw, children might start out with the D position underspecified for phi features relative to adults. This has certain advantages over the analysis in de Villiers and Roeper, in particular in that it captures the patterns of possibility across adult languages. Expletive determiners only occur when they are lexically selected, but English-speaking children clearly allow them in other positions, such as with a heavy verb. Baauw also finds that Spanish children permit expletive determiners in positions

not allowed in adult Spanish (which does have expletive determiners in restricted conditions). That is, Spanish children also permitted wh-extraction from heavy-verb constructions. However, there was no correlation between this and how they treated the constructions in the pronominal co-reference task. de Villiers and Roeper (1995a) actually never attempted a correlational analysis, instead, we did ANOVA analyses. When Baauw does the same, the results are still not significant but they are in the right direction. Baauw has an interesting suggestion for the discrepancy between the English and Spanish results. On his analysis, the island violation is a weak island violation, not a strong one as it necessarily was under Roeper and de Villiers treatment. It is then not too surprising that children persist in violating the weak island effect after they resolve the co-reference problem, given the other weak island results (Philip and de Villiers, 1992).

But the additional difference is attributed to methodology: we used the story/question technique, which attributes competence based on preference differences between minimally different questions. Baauw is right to point out that his Truth Value Judgment Technique allows asking the children whether a certain reading is permitted at all. Our results may indeed have contained two groups of children among the successes: those who obeyed the barrier on principle and those who just had a preference. It remains a puzzle, nevertheless, why such an arbitrary group might show equivalent treatment of the co-reference possibilities in our study. It also removes methodology, however, as an explanation for the original result of a barrier failure, since the result basically replicates using the arguably superior methodology of the Truth Value Judgment Task.

3.4.3 Maturation

No-one has proposed directly that the kinds of errors observed in the wh-barrier work, in particular the medial question answers and false complement error, might relate to maturation. However, maturational accounts of the sort proposed above for the Principle B delay, could be mounted. I end with this discussion because it brings us back full circle to the notion of Modularity in the broadest sense, not in the methodologically imperative sense of the Modularity matching model. Consider again the error in which children answer the wrong thing to:

68) What did she say she bought?

when the character said she bought one thing, but actually bought another. Why might the child answer this way? There are three classes of explanation discussed in de Villiers (1999):

- i. Cognitive immaturity, namely a failure to understand a difference between mental representation and reality.
- ii. Processing limitations, that is, a problem with parsing two clause-sentences.
- iii) Grammatical difficulty in complementation.

The cognitive immaturity argument would be that the child lacks understanding of others' minds, lacks a full "Theory of Mind" so cannot follow a story in which people tell lies, make mistakes, or have false beliefs. In other words, mapping the content is the problem, not the grammar per se. Then, once children do have a full Theory of Mind in the cognitive domain, they will master the sentences in question. There is plenty of evidence that Theory of Mind develops over the course of early childhood. And according to one major position in the field, maturation is what drives Theory of Mind development. On that view, Theory of Mind is yet another module, with its own characteristic inputs, time course and putative brain region (Leslie, 1994; Scholl & Leslie, 1999).

As a variation on the cognitive immaturity argument, one could adapt the solution proposed by Avrutin (1999) and Wexler (1999), in which the child's grammar is complete but inferences about what others know are underdeveloped, and hence the interface in question is not yet adult. But in the Principle B literature, that solution arose after a considerable effort to demonstrate Continuity in Principle B itself, with only minor details of deixis and co-reference left unaccommodated. In the case of complementation, if we do not assume complete grammar, Continuity is at risk. But if one assumes complete grammar, then why is information about others' states of mind intruding into comprehension of complement structures? That would seem to violate Modularity.

The most interesting part of the problem is that the order of development in this case does not seem to be the obvious one, in which the conceptual or inferential delays in development hold back a mature analysis of the sentence. After discussing the alternatives, I conclude (de Villiers 1999) that the mistake is due to an incompleteness in

the grammar of the child, that it is missing a feature in the CP that allows the selection of a clause that is potentially false, i.e. the opposite of factive. In other words, the alternative account entails slow feature-setting on the CP. As in the accounts of the medial question, I contended that the child could not immediately set all the properties pertinent to subcategorization, in particular, the feature set of CP pertinent to complementation.

Once that grammar is achieved, the child becomes capable of answering questions correctly that involve false propositions in their complements, and also producing sentences with false complements:

69) He thought it was a dog

which are surprisingly rare in young children's speech (Bartsch & Wellman, 1995; de Villiers & de Villiers, 1999). But now questions of Modularity come up in an interesting way. In our latest work we have argued that when the child becomes capable of producing sentences with the form and meanings of (69), he also becomes capable of reasoning with such structures. That is, he can entertain thoughts about other peoples' states of belief, about the contents of other minds, that he was not able to do before. In other words, he can answer correctly on so-called "false belief" tasks that require representing the false content of another's beliefs.

Note the twist: instead of mature Theory of Mind being necessary for full comprehension of the sentence, the full grammar of the sentence apparently enables a form of reasoning that did not (or perhaps, could not) exist in the younger child. As Gleitman (p.c.) exclaimed, this could be "the ultimate in syntactic bootstrapping!" Our work with normally developing preschool children suggests support for that direction of effect (de Villiers & Pyers, 1997), and more conclusively, our work with language-delayed oral deaf children (de Villiers & de Villiers, 2000, and in press). Though the cognitive immaturity account would expect language-delayed deaf children to master nonverbal tasks concerning Theory of Mind before the equivalent verbal tasks, in fact both depend on the mastery of complement structures, and hence both are delayed. Thus is the Modularity of Theory of Mind put at risk, in that it depends on certain linguistic prerequisites. But that's someone else's problem. Modularity of language does not preclude the output of the human language faculty from entering into and affecting thought, quite the reverse:

“The faculty of language enters crucially into every aspect of human life, thought and interaction (Chomsky 2000, p. 3).

3.4.4 Summary

The work on wh-questions and barriers also represents an example of principles that the child should know by virtue of having UG from the start. The major findings imply early knowledge of at least the strong island constraints. However, other departures from the adult grammar occur: answering the medial question, lack of weak island effects, and in the false complement mistake. In each case, explanation in terms of parametric variation across languages seems to hold the key, rather than performance error or maturation. The lexically specific nature of complementation is compatible with the argument that at least some parameter setting is a necessarily slow process.

4.0 Conclusion

Paradigms have to have rules and constraints for productive and fine-grained work to get done. Continuity can be seen as one of the strict disciplinarians of the field: stray from this theoretical path at your peril! (It gives a whole new connotation to “Binding Theory”). The other disciplinarian is Modularity, which translates into the methodological imperative: get the experiment right, and don’t sink into the mud of performance excuses. But the rules need periodic re-examination. The most useful criticism will come from researchers who also belong to the paradigm, who understand the foundation and the principles. From that common ground will emerge fruitful discussion and questioning.

In this selective review I wanted to articulate the core features of the generative paradigm, to examine the network of theoretical commitments and forms of argument with a few chosen cases that have formed the model problems for the last several years. The value of such an exercise is in part a synthesis, a review, a recent history to use for educating students. But I also hoped it would illuminate where the cracks are appearing in the paradigm, so I can either decide to slap on more tar and keep it afloat, or jump into a less-seasoned lifeboat with some other crowd (Cognitive grammar? Lexical-functional

grammar? Do Connectionists even have a boat or does a network keep them afloat?) Theoretical work in language acquisition is very, very hard, and needs self-criticism. It is in that spirit that I undertook the review.

It was also to help to sort out what I believe. As I attend conferences and read papers and listen to talks, I hear the kinds of arguments illustrated here, and I am often conflicted. The greatest accolades within the paradigm would be reserved for work that demonstrates early, subtle knowledge, preferably by infants, of just the kinds of principles that UG regards as central, unlearnable, and abstract. My prediction is that such work would hardly receive scrutiny as to its methodology, number of subjects, or conclusions. Indeed, reputations would be made and tenures received. Why don't I cite any? Because it is as rare as hens' teeth, at least in studies of the youngest children. By age four, examples are trivially easy to find, but what is seen before that is more mixed and subject to interpretation. When early competency is elusive, this creates difficult choices of explanation. Here we have discussed three varieties: a) performance excuses, b) parameter setting, and c) maturation. Let me summarize where the difficulties arise.

a) Arguments based on performance have not fared well in general, but it is important to be clear why. If the reasons in each case had just been theoretical, namely, that the resulting model would violate Modularity considerations, it would not have been convincing. But in each of the cases reviewed here, the empirical facts proved more subtle than a performance account would predict, and hence the grammatical accounts have prevailed. Modularity considerations provide the motivation for seeking an alternative explanation, but the arguments need empirical backing.

The issue of how empirical data on performance connect to the principle of Continuity is more troublesome. Suppose children demonstrate subtle knowledge in one performance but not the other. We saw it here in the use of pronouns and reflexives: the failure to observe Principle B is reported in comprehension but not production. So it is in the mastery of sentence complements: apparently good production before complete understanding. Alternatively, experimental procedures testing comprehension could discover some linguistic sensitivity not evident in production. The tendency is to believe the earliest data, and write the other off to a problem in the experiment. That is, we are

biased to sacrifice Modularity on the altar of Continuity. To treat the data fairly, the alternative possibility should at least be acknowledged as a logical possibility, namely that the early correct performance is the artifact, created by some special or limited set of supporting context or scaffolding, and not truly reflective of the child's grammar at that point. Our readiness to see Continuity might lead us to miss important developmental changes.

b) An alternative to early and complete competence, is that the child needs to have certain essential experiences before he can accurately set the parameters of his language. On this account, children's initial grammars might always be compatible with UG, but maybe with the wrong particular grammar. The onus is on such theories to prove that other grammars exist with the setting that children wrongly adopt, and to account for why that setting is adopted, and how it gets reset. It is the account with the most degrees of freedom, given our current state of knowledge about UG, but also the most demanding in terms of providing a semi-learning theory. The fine line such a theory has to walk is this: the role of experience must be enough to reset the parameter, but not so rich as to suggest that the parameter isn't needed. That is why researchers in the generative paradigm have preferred the word "trigger" to "learning experience". That is, if the account is so convincing as to make clear how the parameter gets set, can we still be sure that experience isn't also enough to learn the whole thing?

c) Maturation is a third explanation for the delay in children's competence, one that befits a biological framework. But psychologists are not used to the concept of unfilled time as a causal agent, and it is hard for us to resist the intuition that the time is filled with experiences that might influence the change. But the concept of maturation was invoked exactly to avoid the connotation that experience, i.e. learning, can create the UG principles. Does accident provide us with a natural experiment in which a child grows healthily from age 0 to 4 years with no language input, then is "awakened" to that input at age 4 with an intact cognition and social readiness? Some might argue that profoundly deaf children receiving late cochlear implants provide us with such a test case, though many questions can be raised about its validity. If it is a valid test, the question will be:

will such a child go through the same stages on a delayed timetable, suggesting experience matters, or instantly show the availability of principles posited to require maturation alone?

This review gives but a glimpse of the variety of proposals of researchers working within the constraints of the generative theory of acquisition. The principles of Continuity and Modularity set limits on that theorizing, and are probably worth maintaining. As has been seen, there is still considerable room for alternative positions even when both principles are upheld. Of the two, Continuity ranks higher: it has a longer history as a principle within the paradigm and there are greater costs, in new learnability problems, if it is violated. Modularity *per se* might have to be kept separate from the methodological imperatives that dominate discussion about it.

Given the uncertainties in our state of knowledge of UG and of the language faculty, we should be cautious not to let these limitations become too tight. For example, we still do not have a complete inventory of possible parameters and formal features: linguistic work is expanding our knowledge daily. In short, what seemed like a violation of Continuity yesterday might be discovered not to be tomorrow. Neither do we have a well-worked out theory of the boundaries between syntax and pragmatics, or between pragmatics and real-world knowledge. So where are the edge of the Modules? The incompleteness of our state of knowledge in both these respects should encourage a broader rather than a narrower range of efforts.

BIBLIOGRAPHY

Abdulkarim, L. (2000). Complex wh-questions and universal grammars: New evidence from the acquisition of negative barriers. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Abdulkarim, L. & Roeper, T. (1997). Semantics or syntax for negative islands in language acquisition. In A. Greenhill et al. (Eds.), The Proceedings of the 22nd Annual Boston University Conference on Language Development, Volume 1 (pp.

39-49). Somerville, MA: Cascadilla Press.

Aldridge, M., Borsley, R.D., Clack, S., Creunant, G. & Jones, B.M. (1997). The acquisition of nounphrases in Welsh. In A. Sorace, C. Heycock and R. Shillcock, (Eds.), Proceedings of the GALA 97 conference on language acquisition. Edinburgh, HCRC.

Avrutin, S. (1994). Psycholinguistic investigations in the theory of reference. Doctoral dissertation, MIT.

Avrutin, S. (1999). Development of the Syntax-Discourse Interface. Norwell, MA: Kluwer.

Baauw, S. (2000). Grammatical features and the acquisition of reference. Utrecht: Landelijk Onderzoekschool Taalwetenschap.

Bailyn, J. (1992). LF-movement of anaphors and the acquisition of embedded clauses in Russian. Language Acquisition, 2(4), 307-336.

Bartsch, K. & Wellman, H.M. (1995). Children talk about the mind. New York: Oxford University Press.

Bates, E. & MacWhinney, B. (1989). Functionalism and competition model. In B. MacWhinney and E. Bates (Eds.), The cross-linguistic study of sentence processing (pp. 3-73). Cambridge, UK: Cambridge University Press.

Berman, R.A. (1990). On acquiring an (S)VO language: Subjectless sentences in children's Hebrew. Linguistics, 28, 1135-1166.

Bloom, L. (1970). Language development: Form and function in emerging grammars. Cambridge, MA: MIT Press.

- Bloom, L. (1991). Language development from two to three. New York: Cambridge University Press.
- Bloom, L., Lightbown, P. & Hood, L. (1975). Structure and variation in child language, Monographs of the Society for Research in Child Development, 40(2) serial no.60.
- Bloom, P. (1990). Subjectless sentences in child grammar. Linguistic Inquiry, 21, 491-504.
- Bloom, P. (1999). Theories of word learning: Rationalist alternatives to associationism. In W.C. Ritchie & T.K. Bhatia (Eds.), Handbook of child language acquisition (pp. 249-278). San Diego, CA: Academic Press.
- Borer, H. & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), Parameter setting (pp. 123-172). Dordrecht: Reidel.
- Braine, M. (1973). Three suggestions regarding grammatical analyses of children's language. In C. Ferguson & D. Slobin (Eds.), Studies in child language development. New York: Holt, Rinehart and Winston.
- Brown, R. (1973). A first language: The early stages. Cambridge, MA: Harvard University Press.
- Chien, Y. & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. Language Acquisition, 1(3), 225-295.
- Chierchia, G. (1998). References to kinds across languages. Natural Language Semantics, 6(4), 339-405.

- Chomsky, N. (1965). Aspects of the Theory of Syntax. Cambridge, MA: MIT Press.
- Chomsky, N. (1981). Lectures on Government and Binding. Dordrecht: Foris.
- Chomsky, N. (1982). Some concept and consequences of the theory of government and binding. Cambridge, MA: MIT Press.
- Chomsky, N. (1985). Barriers. Cambridge, MA: MIT Press.
- Chomsky, N. (1992). The minimalist program for linguistic theory. MIT Occasional Papers in Linguistics, 1, Cambridge, MA: MIT Department of Linguistics and Philosophy.
- Chomsky, N. (1995). The minimalist program. Cambridge, MA: MIT Press.
- Chomsky, N. (2000). New horizons in the study of language and mind. New York: Cambridge University Press.
- Cinque, G. (1990). Types of A-dependencies. Cambridge, MA: MIT Press.
- Clahsen, H. (1990). Constraints on parameter setting: A grammatical analysis of some acquisition stages in Child German. Language Acquisition, 1, 361-391.
- Clahsen, H. & Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-Corpus. In J.M. Meisel (Ed.), The acquisition of verb placement (pp. 181-225). Dordrecht: Kluwer.
- Clark, R. (1992). The selection of syntactic knowledge. Language Acquisition, 2, 83-149.
- Clark, R. & Roberts, I. (1993). A computational model of language learnability and

language change. Linguistic Inquiry, 24, 299-345.

Cowie, F. (1999). What's within? Oxford: Oxford University Press.

Crain, S. (1991). Language acquisition in the absence of experience. Behavioral and Brain Sciences, 14, 597-612.

Crain, S. & Lillo-Martin, D. (1999). An introduction to linguistic theory and language acquisition. Oxford, UK: Blackwell.

Crain, S. & Thornton, R. (1991). Recharting the course of language acquisition: Studies in elicited production. In N.A. Krasnegor & D.M. Rumbaugh (Eds.), Biological and behavioral determinants of language development (pp. 321-337). Hillsdale, NJ: Erlbaum.

Crain, S. & Thornton, R. (1998). Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics. Cambridge, MA: MIT Press.

Crain, S., Thornton, R., Boster, C., Conwey, L. Lillo-Martin, D. & Woodams, E. (1996). Quantification without qualification. Language Acquisition, 5, 83-153.

Crain, S. & Wexler, K. (1999). Methodology in the study of language acquisition: A modular approach. In W.C. Ritchie & T.K. Bhatia (Eds.), Handbook of child language acquisition (pp. 387-425). San Diego: Academic Press.

Crisma, P. (1992). Why and how come. Unpublished manuscript, MIT, Cambridge.

de Villiers, J.G. (1991) Why questions? In B. Plunkett and T. Maxfield (eds), The Acquisition of Wh. University of Massachusetts Occasional Papers in Linguistics, 1991.

- de Villiers, J.G. (1992). On the acquisition of functional categories: A general commentary. In J. Meisel (Ed.), The acquisition of verb placement: Functional categories and V2 phenomena in language development (pp. 423-443). Dordrecht: Kluwer.
- de Villiers, J.G. (1999). On acquiring the structural representations for false complements. In B. Hollebrandse (Ed.), New perspectives on language acquisition. Amherst: UMOP.
- de Villiers, J.G., Curran, L., Philip, W. & DeMunn, H. (1998). Acquisition of the quantificational properties of mental predicates. In Proceedings of the 22nd Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.
- de Villiers, J.G., de Villiers, P.A. & Hoban, E. (1994). The central problem of functional categories in the English syntax of oral deaf children. In H. Tager-Flusberg (Ed.), Constraints on language acquisition: Studies of atypical children (pp. 9-47). Hillsdale, NJ: Erlbaum.
- de Villiers, J.G. & de Villiers, P.A. (2000). Linguistic determinism and false belief. In P. Mitchell & K. Riggs (Eds.), Children's reasoning and the mind (pp. 189-226). Hove, UK: Psychology Press.
- de Villiers, J.G. & de Villiers, P.A. (in press). Language for thought: Coming to understand false beliefs. In D. Gentner and S. Goldin-Meadow (Eds.), Advances in the investigation of language and thought. Cambridge, MA: MIT Press.
- de Villiers, J.G. & Pyers, J. (submitted). Complements to Cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding.
- de Villiers, J.G. & Roeper, T. (1991). Introduction: Acquisition of wh-movement. In T.

Maxfield & B. Plunkett (Eds.), University of Massachusetts Occasional Papers. Papers in the acquisition of WH: Proceedings of the UMass Roundtable, May 1990 (pp.1-18). Amherst, MA: University of Massachusetts.

de Villiers, J.G. & Roeper, T. (1993). The emergence of bound variable structures. In E. Reuland & W. Abraham (Eds.), Knowledge and language: Orwell's problem and Plato's problem (pp.105-139). Dordrecht: Kluwer.

de Villiers, J. & Roeper, T. (1995a). Barriers, binding, and acquisition of the DP-NP distinction. Language Acquisition, 4, 73-104.

de Villiers, J.G. & Roeper, T. (1995b). Relative clauses are barriers to wh-movement for young children. Journal of Child Language, 22, 389-404.

de Villiers, J.G. & Roeper, T. (1996). Questions after stories: Supplying context and removing it as a variable. In D. McDaniel, H. Cairns & C. McKee (Eds.), Methodology in child language research (pp. 163-187). Hillsdale, NJ: Erlbaum.

de Villiers, J., Roeper, T., & Vainikka, A. (1990). The acquisition of long-distance rules. In L. Frazier & J. de Villiers (Eds.), Language processing and language acquisition (pp. 257-297). Dordrecht: Kluwer.

Demuth, K. (1994). On the underspecification of functional categories in early grammars. In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 1: Head, projections, and learnability (pp.119-135). Hillsdale, NJ: Erlbaum.

Deprez, V. (1994). Underspecification, functional projections, and parameter setting. In B. Lust, M. Suñer, & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 1: Head, projections, and learnability (pp. 249-272). Hillsdale, NJ: Erlbaum.

- Fanselow, G. & Mahajan, A. (2000). Towards a minimalist theory of wh-expletives, wh-copying, and successive cyclicity. In U. Lutz, G. Müller & A. Von Stechow (Eds.), Wh-scope marking. Amsterdam: Benjamins.
- Fodor, J.A. (1981). Modularity of Mind. Cambridge, MA: MIT Press.
- Fodor, J.D. (1998). Unambiguous triggers. Linguistic Inquiry, 29, 1-36.
- Fukui, N. & Speas, M. (1986). Specifiers and projections. MIT Working Papers in Linguistics, 8, 128-172.
- Gerken, L. (1991). The metrical basis of children's subjectless sentences. Journal of Memory and Language, 33, 19-38.
- Gerken, L. & McIntosh, B. (1993). Interplay of function morphemes and prosody in early language. Developmental Psychology, 29(2), 448-457.
- Gibson, E. & Wexler, K. (1994). Triggers. Linguistic Inquiry, 25, 355-407.
- Gold, E.M. (1967). Language identification in the limit. Information and Control, 10, 447-474.
- Golinkoff, R., Hirsch-Pasek, K., Cauley, K. & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. Journal of Child Language, 14, 23-45.
- Goodluck, H., Foley, M. & Sedivy, J. (1992). Adjunct islands and acquisition. In H. Goodluck & M. Rochemont (Eds.), Island constraints (pp. 181-194). Dordrecht: Kluwer.

- Grimshaw, J. (1994). Minimal projection and clause structure. In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 1: Head, projections, and learnability (pp. 75-84). Hillsdale, NJ: Erlbaum.
- Grimshaw, J. & Rosen, S.T. (1990a). Knowledge and obedience: The developmental status of the binding theory. Linguistic Inquiry, 21(2), 187-222.
- Grimshaw, J. & Rosen, S.T. (1990b). Obeying the binding theory. In L. Frazier & J. de Villiers (Eds.), Language processing and language acquisition (pp. 357-367). Dordrecht: Kluwer.
- Grodzinsky, Y. & Reinhart, T. (1993). The innateness of binding and the development of coreference: A reply to Grimshaw and Rosen. Linguistic Inquiry, 24(1), 69-103.
- Guasti, T. (1994). Verb syntax in Italian child grammar: Finite and nonfinite verbs. Language Acquisition, 3(1), 1-40.
- Guilfoyle, E. & Noonan, M. (1988). Functional categories and language acquisition. Paper presented at the 13th Annual Boston University Conference on Language Development.
- Hamburger, H. & Crain, S. (1982). Relative acquisition. In S. Kuczaj (Ed.), Language development (1). Hillsdale, NJ: Erlbaum.
- Hirsch-Pasek, K. & Golinkoff, R. (1997). The origins of grammar. Cambridge, MA: Bradford Books.
- Hoekstra, T. & Hyams, N. (1994). Missing heads in child language. In C. Koster and F. Wijnen (Eds.), Proceedings of the GALA 95. Center for Language and Cognition, Groningen.

- Hohle, B. & Weissenborn, J. (1998). Sensitivity to closed class items in preverbal children. In A. Greenhill, M. Highes, H. Littlefield & H. Walsh (Eds.), Proceedings of the 22nd Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.
- Hohle, B. & Weissenborn, J. (2000). The origins of syntactic knowledge: Recognition of determiners in one-year-old German children. In C. Howell, S. Fish & T. Keith-Lucas (Eds.), Proceedings of the 24th Annual Boston University Conference on Language Development, Volume 2. Somerville, MA: Cascadilla Press.
- Hyams, N. (1986). Language acquisition and the theory of parameters. Dordrecht: Reidel.
- Hyams, N. (1987). The theory of parameters and syntactic development. In T. Roeper & E. Williams (Eds.), Parameter Setting. Dordrecht: Reidel.
- Hyams, N. (1991). A reanalysis of null subjects in child language. In J. Weissenborn, H. Goodluck & T. Roeper (Eds.), Theoretical issues in language acquisition: Continuity change in development. Hillsdale, NJ: Erlbaum.
- Hyams, N. (1992). Morphosyntactic development in Italian and its relevance to parameter-setting models: Comments on the paper by Pizzuto and Casselli. Journal of Child Language, 19, 695-709.
- Hyams, N. (1994). Commentary: Null subjects in child language and the implications of cross-linguistic variation. In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 2: Binding, dependencies, and learnability (pp. 287-300). Hillsdale, NJ: Erlbaum.
- Hyams, N. (2000). Now you hear it, now you don't: The nature of optionality in early

child grammar. Plenary address, Boston University Conference on Language Development, November.

Hyams, N. & Sigurjonsdottir, S. (1990). The development of long distance anaphora: A cross-linguistic comparison with special reference to Icelandic. Language Acquisition, 1(1), 57-94.

Hyams, N. & Wexler, K. (1993). On the grammatical basis of null subjects in child language. Linguistic Inquiry, 24, 421-459.

Jackendoff, R. (1977). X' Syntax. Cambridge, MA: MIT Press.

Jacobowicz, C. (1984). On markedness and binding principles. Proceedings of the North East Linguistic Society, 14, 154-182.

Jaeggli, O. & Hyams, N. (1988). Morphological uniformity and the setting of the null subject parameter. Proceedings of the Northeastern Linguistic Society, 18.

Katz, N., Baker, E. & Macnamara, J. (1974). What's in a name? A study of how children learn common nouns and proper names. Child Development, 45, 469-473.

Kohl, K. (1999). An analysis of finite parameter learning in linguistic spaces. Unpublished thesis, MIT, Cambridge, MA.

Kuhn, T.S. (1963). The Structure of Scientific Revolutions, 2nd edition. Chicago, IL.

Lasnik, H. & Stowell, T. (1991). Weakest cross-over. Linguistic Inquiry, 2, 687-720.

Lebeaux, D. (1988). Language acquisition and the form of the grammar. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

- Leslie, A.M. (1994). *Pretending and believing: Issues in the theory of ToMM*. Cognition, 50, 211-238.
- Lillo-Martin, D. (1991). Universal grammar and American Sign Language: Setting the null argument parameters. Dordrecht: Kluwer.
- Lillo-Martin, D. (1999). Modality effects and modularity in language acquisition: The acquisition of American Sign Language. In W. Ritchie & T.K. Bhatia (Eds.), Handbook of child language acquisition (pp.531-568). San Diego, CA: Academic Press.
- Longobardi, G. (1994). Reference and proper names: A theory of N-Movement in syntax and logical form. Linguistic Inquiry, 25, 609-665.
- Lust, B. (1986). Studies in the acquisition of anaphora: Vol. 1. Defining the constraints. Dordrecht: Reidel.
- Lust, B. (1994). Functional projection of CP and phrase structure parameterization: An argument for the strong continuity hypothesis. In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 1: Head, projections, and learnability. (pp. 85-118). Hillsdale, NJ: Erlbaum.
- Lust, B. (1999). Universal grammar: The strong continuity hypothesis in first language acquisition. In W.C. Ritchie & T.K. Bhatia (Eds.), Handbook of child language acquisition (pp. 111-156). San Diego, CA: Academic Press.
- Lust, B., Flynn, S., Foley, C. & Chien, Y-C. (1999). How do we know what children know? Problems and advances in establishing scientific methods for the study of language acquisition and linguistic theory. In W.C. Ritchie & T.K. Bhatia (Eds.), Handbook of child language acquisition (pp. 427-456). San Diego, CA: Academic Press.

- Macnamara, J. (1982). Names for things: A study of human learning. Cambridge, MA: MIT Press.
- MacWhinney, B. (1999). The emergence of language. Mahwah, N.J.: Erlbaum.
- MacWhinney, B. & Snow, C. (1985). The Child Language Data Exchange System. Journal of Child Language, 12, 271-296.
- Maxfield, T. & Plunkett, B. (1991). Papers in the acquisition of WH. UMOP Special Edition, University of Massachusetts, Amherst.
- Mazuka, R. & Lust, B. (1994). When is an anaphor not an anaphor? In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Vol. 2: Binding, dependencies, and learnability (pp. 145-175). Hillsdale, NJ: Erlbaum.
- McDaniel, D. (1989). Partial and multiple Wh-movement. Natural Language and Linguistic Theory, 7, 565-604.
- McDaniel, D. & Maxfield, T. (1992). Principle B and contrastive stress. Language Acquisition, 2(4), 337-358.
- McDaniel, D., Chiu, B., & Maxfield, T. (1995). Parameters for wh-movement types: Evidence from child language. Natural Language and Linguistic Theory, 13, 709-753.
- McKee, C. (1992). A comparison of pronouns and anaphors in Italian and English acquisition. Language Acquisition, 2, 21-54.
- McKee, C. (1994). What you see isn't always what you get. In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic

perspectives, Volume 1: Head, projections, and learnability (pp. 201-212). Hillsdale, NJ: Erlbaum.

McKee, C., Nichol, J. & McDaniel, D. (1993). Children's application of binding during sentence processing. Language and Cognitive Processes, 8(3) 265-290.

McNeill, D. (1970). The development of language. In Paul Mussen (Ed.), Carmichael's handbook of child psychology, Vol 1 . New York: Wiley.

Meisel, J. & Muller, N. (1992). Finiteness and verb placement in early child grammars: Evidence from the simultaneous acquisition of German and French in bilinguals. In J. Meisel (Ed.), The acquisition of verb placement: Functional categories and V2 phenomena in language development. Dordrecht:Kluwer.

Nishigauchi, T. & Roeper, T. (1997). Deductive parameters and the growth of empty categories. In T. Roeper & E. Williams (Eds.), Parameter setting (pp. 91-122). Dordrecht: Reidel.

Otsu , Y. (1981). Universal grammar and syntactic development in children: Toward a theory of syntactic development. Unpublished doctoral dissertation, MIT.

Philip, W. & Coopmans, P. (1996). The double Dutch delay of Principle B effect. In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (Eds.), Proceedings of the 20th Annual Boston University Conference on Language Development . Somerville, MA: Cascadilla Press.

Philip, W. & de Villiers, J.G. (1992). Monotonicity and the acquisition of weak islands. In E. Clark (Ed.), Proceedings of the Twenty-fourth Stanford Child Language Conference (pp. 99-111). Stanford University CSLI.

Philip, W. & de Villiers, J.G. (1993). The relation between logical and syntactic

- development. Paper presented at symposium at Sixth International Congress on Child Language, Trieste, Italy.
- Pinney, M. (1981). Syntactic constraints and the acquisition of embedded sentences. Unpublished doctoral dissertation, University of Massachusetts, Amherst.
- Pierce, A. (1992). Language acquisition and syntactic theory: A comparative analysis of French and English child grammars. Dordrecht: Kluwer.
- Pinker, S. (1984). Language learnability and language development. Cambridge, MA: Harvard University Press.
- Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. Cognition, 28, 73-193.
- Pizzuto, E. & Casselli, M.C. (1992). The acquisition of Italian morphology: Implication for models of language development. Journal of Child Language, 19, 491-557.
- Platzack, C. (1990). A grammar without functional categories: A syntactic study of early Swedish child language. Nordic Journal of Linguistics, 13, 107-126.
- Poepfel, D. & Wexler, K. (1993). The full competence hypothesis. Language, 69, 1-33.
- Pollock, J-Y. (1989). Verb movement, universal grammar and the structure of IP. Linguistic Inquiry, 20, 365-424.
- Radford, A. (1988). Small children's small clauses. Transactions of the Philological Society, 86, 1-46.
- Radford, A. (1990a). Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English. Oxford, UK: Blackwell.
- Radford, A. (1990b). The syntax of nominal arguments in early child English. Language

Acquisition, 1(3), 195-223.

- Radford, A. (1994). Tense and agreement variability in child grammars of English. In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 1: Head, projections, and learnability (pp. 135-158). Hillsdale, NJ: Erlbaum.
- Radford, A. (1995). Children- architects or brickies? In D. Maclaughlin & S. McEwen (Eds.), Proceedings of the BUCLD19 (pp.1-19). Somerville, MA: Cascadilla Press.
- Radford, A., Atkinson, M., Britain, D., Clahsen, H. & Spencer, A. (1999). Linguistics: An introduction. Cambridge, UK: Cambridge University Press.
- Randall, J. (1990). The catapult hypothesis: An approach to unlearning. In J. Weissenborn, H. Goodluck & T. Roeper (Eds.), Theoretical issues in language acquisition: continuity change in development. Hillsdale, NJ: Erlbaum.
- Reinhart, T. (1983). Anaphora and semantic interpretation. London, UK: Crumb Helm.
- Reinhart, T. & Reuland, E. (1993). Reflexivity. Linguistic Inquiry, 24, 657-720.
- Rizzi, L. (1982). Issues in Italian Syntax. Dordrecht: Foris.
- Rizzi, L. (1990). Relativized minimality. Cambridge, MA: MIT Press.
- Rizzi, L. (1994). Early null subjects and root null subjects. In B. Lust, G. Hermon & J. Kornfilt (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Vol. 2: Binding, dependencies, and learnability. (pp. 249-272). Hillsdale, NJ: Erlbaum.
- Roeper, T. (1992). From the initial state to V2: Acquisition principles in action. In J.M.

- Meisel (Ed.), The acquisition of verb placement (pp. 333-371). Dordrecht: Kluwer.
- Roeper, T. (1996) The role of Merger Theory and formal features in acquisition. In H. Clahsen (Ed.), Generative perspectives on language acquisition. Amsterdam: John Benjamins.
- Roeper, T. (1998). Finding fundamental operations in language acquisition: Formal features as triggers. In B. Hollebrandse (Ed.), New perspectives on language acquisition (pp. 197-214). Amherst, MA: UMOP/GLSA.
- Roeper, T. (1999). Universal bilingualism. Bilingualism: Language and Cognition, 2, **pp.**
- Roeper, T., & de Villiers, J. (1992). Ordered decisions in the acquisition of Wh-questions. In J. Weissenborn, H. Goodluck & T. Roeper (Eds.), Theoretical issues in language acquisition: Continuity and change in development (pp. 191-236). Hillsdale, NJ: Erlbaum.
- Roeper, T., & de Villiers, J. (1994). Lexical links in the WH-chain. In B. Lust, G. Hermon & J. Kornfilt (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 2: Binding, dependencies, and learnability (pp. 357-390). Hillsdale, NJ: Erlbaum.
- Roeper, T. & Rohrbacher, B. (1994). Null subjects in early child English and the theory of economy of projection. Unpublished manuscript, University of Massachusetts, Amherst and University of Pennsylvania, Philadelphia.
- Roeper, T. & Wiessenborn, J. (1990). How to make parameters work. In L. Frazier & J.G. de Villiers (Eds.), Language processing and language acquisition (pp.147-162). Dordrecht: Kluwer.
- Roeper, T. & Williams, E. (1987). Parameter Setting. Boston, MA: Reidel.

- Ross, J.R. (1967). Constraints on variables in syntax. Doctoral dissertation, MIT.
Published as Infinite syntax!, Norwood, NJ: Ablex (1986).
- Sabbagh, M. & Gelman, S. (2000). Buzzsaws and blueprints: What children need or don't need to learn language. Journal of Child Language, 27, 715-726.
- Sano, T. & Hyams, N. (1994). Agreement, finiteness and the development of null arguments. In Proceedings of North Eastern Linguistics Society, 24.
- Scholl, B. & Leslie, A. (1999). Modularity, development and 'Theory of Mind'. Mind and Language, 14, 131-153.
- Schaeffer, J. (2000). The modularity of grammar and pragmatics: Evidence from SLI. Paper presented at the Boston University Conference on Language Development, November.
- Shady, M.E. (1996). Infants' sensitivity to function morphemes. Unpublished doctoral dissertation, State University of New York at Buffalo.
- Shady, M., Jusczyk, P. & Gerken, L. (1998). Infants' sensitivity to function morphemes. In E. Hughes, M. Hughes & A. Greenhill (Eds.), Proceedings of the 21st Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.
- Shipley, E., Smith, C. & Gleitman, L. (1969). A study in the acquisition of language: Free responses to commands. Language, 45, 322-342.
- Sigurdjondottir, S. (1992). Binding in Icelandic: Evidence from language acquisition. Unpublished doctoral dissertation, University of California.

- Speas, M. (1990). Phrase structure in natural language. Dordrecht: Kluwer.
- Speas, M. (1994). Null arguments in a theory of economy of projection. In E. Benedicto & J. Runner (Eds.), Functional Projections. Amherst, MA: UMOP, 17.
- Stowell, T. (1981). Origins of phrase structure. Cambridge, MA: MIT Press.
- Stowell, T. (1991). Determiners in NP and DP. In K. Leffel & D. Bouchard (Eds.), Views on phrase structure. Dordrecht: Kluwer.
- Szabolsci, A. & Zwarts, F. (1992). Unbounded dependencies and the algebraic semantics. Lecture notes of the Third European Summer School in Logic, Language and Information. Saarbrücken, August 1991.
- Thornton, R. (1991). Adventures in long distance moving: The acquisition of complex wh-questions. Unpublished doctoral dissertation, University of Connecticut.
- Vainikka, A. & Levy, Y. (1999). Empty subjects in Finnish and Hebrew. Natural Language and Linguistic Theory, 17, 613-671.
- Valian, V. (1990). Null subjects: A problem for parameter-setting models of language acquisition. Cognition, 35, 105-122.
- Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. Cognition, 40, 21-82.
- Valian, V. (1992). Categories of first syntax. *be, be-ing, and nothingness*. In J. Meisel (Ed.), The acquisition of verb placement: Functional categories and V2 phenomena in language acquisition. Dordrecht: Kluwer.
- Valian, V. (1994). Children's postulation of null subjects: Parameter setting and language

acquisition. In B. Lust, G. Hermon & J. Kornfilt (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 2: Binding, dependencies, and learnability (pp. 273-300). Hillsdale, NJ: Erlbaum.

Valian, V., Hoeffner, J. & Aubrey, S. (1996). Young children's imitation of sentence subjects: evidence of processing limitations. Developmental Psychology, 23, 153-164.

Van Valin, R.D. Jr., (1998). The acquisition of WH-questions and the mechanisms of language acquisition. In M. Tomasello (Ed.), The New Psychology of Language: Cognitive and functional approaches to language structure (pp. 221-249). Hillsdale, NJ: Erlbaum.

Verrips, M. & Weissenborn, J. (1992). Route to verb placement in early German and French: The independence of finiteness and agreement. In J.M. Meisel (Ed.), The acquisition of verb placement (pp. 283-333). Dordrecht: Kluwer.

Wang, Q., Lillo-Martin, D, Best, C.T. & Levitt, A. (1992). Null subjects and objects in the acquisition of Chinese. Language Acquisition, 2, 221-254.

Weissenborn, J. (1994). Constraining the child's grammar: Local well-formedness in the development of verb movement in German and French. In B. Lust, M. Suñer & J. Whitman (Eds.), Syntactic theory and first language acquisition: Cross-linguistic perspectives, Volume 2: Head, projections, and learnability (pp. 215-248). Hillsdale, NJ: Erlbaum.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot & N. Hornstein (Eds.), Verb Movement (pp. 305-382). Cambridge, UK: Cambridge University Press.

Wexler, K. (1996). The development of inflection in a biologically based theory of language acquisition. In M. Rice (Ed.), Toward a genetics of language.(pp. 113-144)

Hillsdale, NJ: Erlbaum.

- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. Lingua, 106 (1-4), 23-79.
- Wexler, K. (1999). Maturation and growth of grammar. In W. Ritchie & T.K.Bhatia (Eds.), Handbook of child language acquisition (pp. 55-109). San Diego, CA: Academic Press.
- Wexler, K. & Chien, Y-C. (1986). Development of lexical anaphors and pronouns. Papers and Reports on Child Language Development, Vol 24. Stanford University.
- Wexler, K., & Manzini, M. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds.), Parameter setting (pp. 41-76). Dordrecht: Reidel.
- Yang, C.D. (2000). Knowledge and learning in natural language. Unpublished doctoral dissertation, MIT.
- Yang, C.D. & Guttman, S. (1999). Language learning via Martingales. Paper presented at The Sixth Conference of the Mathematics of Language, Orlando, FL.
- Zubizarreta, M.L. & Vergnaud, J.R. (1992) The definite determiner and the inalienable construction in French and English. Linguistic Inquiry, 23, 595-652.