PARAMETERS, LEARNABILITY, AND THE NULL SUBJECT PARAMETER

MARTHA SCHULTE-NAFEH & PAOLA DUSSIAS
University of Arizona

Parameters were originally proposed within the Principles and Parameters model of linguistic analysis (Chomsky, 1981) as a means of explaining the relative ease and rapidity with which children construct grammars for their target languages. These parameters were seen as linking together seemingly unrelated surface syntactic features so that once "set" on the basis of simple input data, readily available to the child, a number of syntactic features would be "learned" simultaneously. In this paper, we consider the evolution of the concept of the parameter and show how learnability considerations and cross-linguistic variation have led to a very different notion of parameters based on the goal of preventing overgeneralization, in contrast to the original goal of preventing undergeneralization. Turning to the specific case of the null subject parameter, we evaluate various proposals (Hyams, 1986, 1987a; Bloom, 1990, 1993; Valian, 1990b) for what the initial setting of that parameter is, what the triggering data is for setting it, and what markedness relation exists between different values of the parameter. It becomes clear that the original conceptualization of a parameter faces serious empirical challenges when one attempts to specify precisely how it might work in particular cases.

1. Introduction

In this paper we consider some constructs within the theories of learnability and markedness associated with the broader framework of theoretical syntax referred to as Government and Binding, or alternatively, the Principles and Parameters theory. These constructs will be considered first in general terms. Then we will look at how they have been modified as a result of their application to specific situations, with particular focus on the case of the null subject parameter.

2. The Principles and Parameters Framework

One of the central goals of the Principles and Parameters theory has been to characterize grammars for natural languages in such a way that they have "explanatory adequacy." That is, they provide or are consistent with an explanation of how children come to construct adult-like grammars based on what we know about the input to this learning process. Under normal circumstances, all children will learn the native language of whatever community they are born into. In doing so they are provided with an abundance of positive evidence about what is a possible grammatical string in the target language but they do not receive negative evidence in that they are not explicitly told what strings are ungrammatical, nor are they corrected for producing ungrammatical strings. Brown and Hanlon (1970) showed that signs of approval or disapproval that children receive from caregivers are not a function of the grammatical correctness of their utterance but rather are a response to the truth of the proposition expressed. "Explicit approval or disapproval of either syntax or morphology is extremely rare in our records and so seems not to be the force propelling the child from immature to mature forms" (p. 48).[1]

The fact that children come to construct a grammar of their target language based on impoverished data has been called Plato's problem (Chomsky 1986). The fact that children do not receive negative evidence is one aspect of this impoverishment. A central concern in modern linguistic theorizing has been to explain how we come to know so much about the grammar of our

language as adults given that the data from which we infer this grammar is impoverished. The answer to this question offered within the Principles and Parameters theoretical framework is that children are endowed with an innate language acquisition device (LAD) that encodes a universal grammar (UG) which underlies their ability to learn and perhaps to produce and comprehend their spoken language. This universal grammar is of a general enough nature that it can be used by children to acquire whatever specific language they are exposed to, operating on the input provided by members of their community speaking around and to them. The basic notion is that while languages differ, they do so in finite ways. There is not an infinite number of ways in which a human language can encode meaning, and the LAD constrains the hypotheses which children will make about the grammar of the language they are acquiring.

In broad terms, UG is thought to involve a set of general principles accounting for the universal aspects of human languages, as well as parameters, which encode the possible variations of the basic universal features. Chomsky describes these parameters as "genetically permitted variations that exist as options in the universal grammar" (Gliedman, 1985, p. 372). The variations are encoded in terms of "values" of the parameters which must be "set" by the learner based on experience with the target language. The values may be binary, having a (+) or (-) value, or they may have multiple values. Once the parameters have been set by the learner, the "core grammar" has been identified. It has also been proposed that the settings of the parameters of the core grammar, in some cases, involve a markedness hierarchy, with one value being less marked than another. The initial parameter setting "hypothesized" by the child is the least marked option.

The parameters, as initially conceptualized, were thought to relate a number of different surface syntactic features of a language, in that these features where expressed in terms of just one parameter. A case in point is the null subject parameter, which was purported to relate 1) the possibility of phonologically null subjects, 2) free subject-verb inversion, and 3) that-trace effects. In this way, "complex patterns of variation are reduced to minimal differences in the parametric choices" (Rizzi, 1982, p. 117). The following quote from Pinker (1990) expresses succinctly this interactional view of parameters:

> The reason this difference [whether or not a given language allows the speaker to omit the subject in a tensed sentence with an inflected verb] is thought of as a "parameter" rather than an isolated fact is that it corresponds not only to the presence or absence of overt subjects but also to a variety of more subtle linguistic facts that are all present in languages with null subjects and absent in languages that require the subject to be overt....Thus, the rules of a grammar interact tightly; a single change will give **a series of cascading effects throughout the grammar** [emphasis added]. On this view, the child only has to set these parameters on the basis of parental input, and the full richness of grammar will ensue when those parametrized rules interact with one another and with universal principles. The parameter setting view can help explain the universality and rapidity of language acquisition: when the child learns one fact about her language, she can deduce that other facts are also true of it without having to learn them one by one. (p. 230)

3. Formal Learning Theory

Recall that one of the central goals of the Principles and Parameters theory has been to provide an explanation of how children acquire language. In seeking such an explanation, researchers have drawn on work in a branch of theoretical computer science called "the mathematical theory of automata," or more simply, learnability theory. The theory is meant to account for all kinds of learning but we will describe it in terms of its application to language learning only.

The theory posits four basic components of a language learning "system": 1) A class of languages, one of which is the target language which is to be learned. 2) An environment. Considered here is the nature of the information about the target language that is available to the learner and its presentation.
3) A learning strategy. The learner, using information in the environment, "hypothesizes" about the target language. Investigated here is the algorithm that creates the hypotheses, or the grammar-

forming mechanisms used by the learner. 4) A success criterion. A specification of what constitutes successful language learning is given. Presumably this involves some level of approximation to adult grammars in the target language.

Specification of any three of the four components places logical restrictions on what the fourth can be. Formal learning theorists attempt to identify which classes of languages are learnable, under what conditions of information presentation to a learner using particular learning strategies. Among the most influential work in this field of formal language learning, in terms of application to natural language learning, is that of Gold (1967) and Angluin (1978, 1980). The theorems developed in this field are of a highly technical nature and we will not consider them here in any detail. Suffice it to say that from these theorems, a very central idea in current theorizing about natural language learnability has evolved: the Subset Principle.

The Subset Principle is intended to provide a solution to the problem of overgeneralization which is possible in certain language learning situations. Before considering this solution, we will spell out the conditions under which the problem arises.

3.1. The Environment and Learning Strategies

As mentioned at the outset, research has shown that children do not receive explicit negative evidence about their target grammars. That is, they are not instructed as to which sentences are ungrammatical. They receive only positive evidence, that is, examples of grammatical strings. In formal learning theory this is referred to as presentation of a text. When the presentation does include negative evidence in the form of examples of ungrammatical sentences, which are designated as such, it is said to be presentation from an informant.

It might seem that these two kinds of evidence, positive and negative, exhaust the possibilities for kinds of information that could theoretically be available to a first language learner. However, a third possibility has been considered and has been called "indirect" or "implicit" negative evidence. The idea is that, under certain circumstances, the absence of certain strings in the text can serve as evidence that such strings are not grammatical in the target language. In fact, indirect negative evidence does not define a different environment (in the formal sense) for learning as we are still dealing with a text presentation of the strings of the language. Rather, it is a learning strategy that the learner is postulated to have available when constructing a grammar based on text presentation. That is, the learner is thought to be able to use the fact of non-occurrence of certain strings in the text as evidence that they cannot occur.

The possibility that learners use this kind of indirect negative evidence in grammar construction was entertained in early formulations of the Principles and Parameters theory. Chomsky (1981) states:

> A not unreasonable acquisition system can be devised with the operative principle that if certain structures or rules fail to be exemplified in relatively simple expressions, **where they would be expected to be found** [emphasis added], then a (possibly marked) option is selected excluding them in the grammar, so that a kind of "negative evidence" can be available even without corrections, adverse reactions, etc. (p. 90)

However, explicating the phrase "where they would be expected to be found" has proven difficult. If children do have a learning procedure which involves the use of indirect negative evidence, the task of specifying exactly how and when they use it is logically a very difficult one. In principle, there are an infinite number of strings that the child doesn't hear; therefore, allowing for such a learning procedure greatly broadens the child's "hypothesis space." The question becomes, which of those infinite number of absent strings are relevant to the learner? Because of these problems, researchers working within the Principles and Parameters approach have preferred to posit other, more deterministic learning mechanisms and principles that can account for successful acquisition from positive evidence only (Berwick, 1985; Lasnik, 1989).

Nonetheless, while justified on logical grounds, the requirement of learning from positive evidence only is, in the final analysis, an assumption within this theoretical framework and not a fact about learning strategies available to children learning their first language. First language

learners certainly are not provided with direct negative evidence, but whether or not they do in some way use indirect or implicit negative evidence remains an open question.[2] In the remainder of this paper, we will use the term negative evidence to refer to presentation by an informant, including examples of ungrammatical strings, and indirect negative evidence to refer to the learning procedure that operates on the absence of certain strings in a text presentation.

3.2. The Class of Languages: The Subset Condition/Property

Given that alternative hypotheses about the grammar of the target language generate different strings of grammatical sentences, we can consider what possible set-theoretic relationships might arise between these languages. We will speak of these hypotheses in terms of different values of parameters since parameter setting is our focus here. If a parameter p has two values: i and j, then we can designate the languages (strings of grammatical sentences) generated by each value of the parameter as $L(p(i))$ and $L(p(j))$, respectively. They may exhibit any of the following relationships to each other:

(1) a)   b)   c)   d)
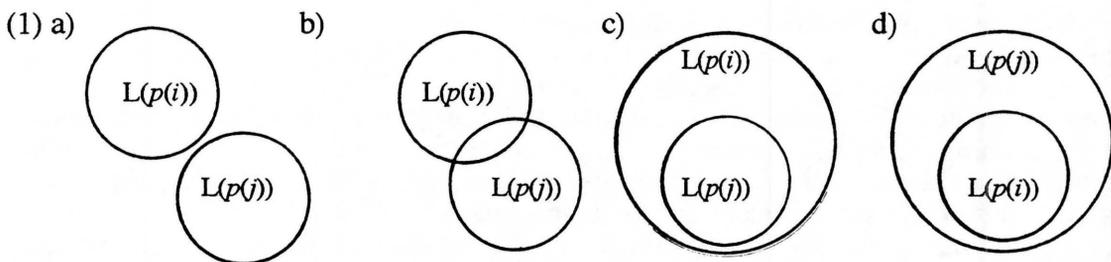
In (a), the sets of grammatical strings generated by each value of the parameter are disjoint. If the learner hypothesizes either $L(p(i))$ or $L(p(j))$ as the target language, and that hypothesis is incorrect, positive evidence will be available to tell the learner that the guess was incorrect and what the correct setting is. That positive evidence will consist of strings in the environment which are not consistent with the hypothesized setting of the parameter. Likewise, in the situation depicted in (b), positive evidence will be available to the learner to correct an incorrect hypothesis about the parameter value in the target language. As long as there is some set of strings in $L(p(i))$ but not in $L(p(j))$, and some set of strings in $L(p(j))$ but not in $L(p(i))$, the learner will have positive evidence of having made an incorrect hypothesis about the target language, whatever the initial hypothesis may be.

However, in the situations depicted in (c) and (d), where either $L(p(j))$ is a subset of $L(p(i))$ or vice-versa, there exists the problem of overgeneralization which cannot be corrected by positive evidence. Specifically, if the learner guesses the superset language, and the target language is the subset language, the environment will not offer positive evidence of the mistake. Since all the strings in the subset are also in the superset, the environment will always be consistent with the hypothesized superset, although it is not the target language. The learner would be producing strings that are not grammatical in the target, but without negative evidence will not be able to correct this situation. In discussing this problem in terms of the weakness of text presentation (positive evidence only) of the target language, Gold (1967) states that "the problem with text is that, if you guess too large a language, the text will never tell you that you are wrong" (p. 461).

So, when languages generated by alternative values of a parameter stand in a subset relation to one another and the learner has only positive evidence on which to base the induction procedure, there is always the potential of an intractable learning problem. If the learner hypothesizes the larger, superset language, there will not be positive evidence of the faulty hypothesis, and so the learner cannot recover from it. Of course if the subset language is hypothesized, there is no problem as there will be positive evidence that the hypothesis is incorrect in the event that it is; that is, if the target language is in fact the superset. There will be sentences in the data which are not consistent with the subset grammar, and they will trigger a switch to the superset grammar.

When the languages generated by two values of a parameter are in a subset/superset relation they are said to exhibit the *Subset Property* (Berwick, 1985). It has been proposed that in such a situation, there is a specific learning strategy available to the learner in terms of a ranking of the parameters. This has been called the *Subset Principle* and is "the key condition guaranteeing identifiability from positive evidence" when languages exhibit the *Subset Property* (p. 235). The Subset Principle essentially establishes a markedness hierarchy for parameter settings in cases where alternative parameter settings yield languages in a subset relation to one another. The unmarked parameter setting, which the learner will initially adopt, is the one which generates the smaller, subset language. This hypothesis will be abandoned in favor of a more marked setting, which generates a larger superset language, only if positive evidence indicates that this more marked setting is the correct one.

In terms of the null subject parameter, considering only the absence or presence of overt subject pronouns, the (-) value of the parameter would be the unmarked setting since it generates strings with overt referential subject pronouns in tensed sentences, while the (+) value generates these strings as well as ones with null subject pronouns. The prediction then would be that children would initially assume a (-) setting for the null subject parameter. They would assume, unless given positive evidence to the contrary, that they cannot omit referential subject pronouns. As we will see below, a question which arises repeatedly in investigation of the pro-drop parameter in first language acquisition is what surface syntactic features are associated with the parameter and whether or not the parameter settings do in fact generate languages (strings of grammatical sentences) that stand in a subset relation to one another.

In his initial formulation of the Subset Principle, Berwick (1985) suggests that the Principle, in conjunction with formulations in UG which lead to the conditions for its application, can replace the need to postulate indirect negative evidence as a learning mechanism. He gives a number of examples from the literature where this is the case.

4. The Lexicalization of Parameters

The formulation of the Subset Principle has had a substantial impact on the conceptualization of parameters within the theory. In a very influential paper, Wexler and Manzini (1987; henceforth W&M) applied the notions of the Subset Condition and the Subset Principle to an investigation of the cross-linguistic instantiation of the governing categories relevant to the Binding theory, a sub-theory within the Government and Binding framework. The Binding Theory deals with the specification of the governing category (syntactic domain) which is operative in the following two basic conditions on well-formedness:

Condition A: An anaphor must be bound in its governing category.
Condition B: A pronoun must be free in its governing category.

As originally conceptualized, the governing category (GC) was thought to be the parameter of Binding theory. That is, every language would exhibit a syntactic domain within which anaphors must be bound and pronouns must be free. This is the universal principle. The parameter was the precise specification of that domain. W&M, based on extensive cross-linguistic data, developed the following five-value parameter necessary to account for observed variation with regard to the governing category:

g is the governing category for a iff
g is the minimal category which contains a and
a. has a subject, or
b. has an INFL, or
c. has a TNS , or
d. has an indicative TNS, or
e. has a root TNS

They go on to show that the parameter setting for anaphors must be independent of that for pronouns, as there are languages (Icelandic is the specific case) in which the GC is not the same for anaphors as it is for pronouns. Furthermore, they show that the markedness hierarchy, based on the Subset Principle, for the GC of anaphors is the opposite of that for pronouns. Specifically, as governing categories for anaphors, the syntactic domains listed above stand in a subset relation to each other as follows: $a \subseteq b \subseteq c \subseteq d \subseteq e$. For pronouns, the relation is: $e \subseteq d \subseteq c \subseteq b \subseteq a$. So, they conclude that if the Subset Principle is the basis for parameter setting, the parameter must be set independently for pronouns and anaphors.

Moreover, they point out that in some languages the governing category may vary for different anaphors in a language, so that the parameter value, rather than being a property of the language as a whole, may indeed need to be associated with particular lexical items in the language, in this case particular anaphors. They formulate the following *Lexical Parameterization Hypothesis:*

> Values of a parameter are associated not with particular languages, but with particular lexical items in a language. (p. 55)

W&M also propose the need for a second parameter in Binding Theory. This second parameter is necessary to account for the observed cross-linguistic variation with regard to what kind of noun phrases can serve as proper antecedents to bind anaphors or from which pronouns must be free. In this case the relevant data is Japanese, in which only subjects can serve as proper antecedents. The parameter they propose is formulated as follows:

> Proper Antecedent Parameter:
> A proper antecedent is:
> a. a subject or
> b. any element

Again, a subset relation holds in opposite directions for anaphors and pronouns. For anaphors, the (a) setting yields strings in a subset relation to those generated by the (b) setting. Looking at pronouns we find that the (b) setting generates the subset stings relative to the (a) setting.

W&M show that in order for the subset condition to hold--that is, to ensure a subset relation between values of a parameter--the parameter for the governing category and that for the proper antecedent must be set independently. If the two parameters are set simultaneously, then the subset condition is not met in all cases. You could get parameter values that would yield strings of sentences in an overlapping relation (as in (1b), above) and hence the Subset Principle could not apply.

In essence, W&M here are defining a parameter as necessarily generating languages in a subset relation to one another. In their proposal, the grammar, in the parameters made available, basically guarantees that the problematic subset relation will arise as there is a learning function--the Subset Principle--which solves the problem. This version of the Subset Condition is much stronger and qualitatively different from the Subset Condition proposed by Berwick (1985). For Berwick, the Subset Condition was descriptive of a situation that might arise between the languages generated by different values of a parameter. W&M suggest that the Subset Condition is a condition on what can be a parameter in the theory:

> Let us assume, then, that the theory of learnability indeed includes a restriction to the effect that the languages generated by two values of a parameter are a subset one of the other, for every given parameter and every two values of it. This restriction we can formulate...and refer to as the *Subset Condition:*

> For every parameter p and every two values i, j of p, the languages generated under the two values of the parameter are one a subset of the other, that is, $L(p(i)) \subseteq L(p(j))$ or $L(p(j)) \subseteq L(p(i))$. (p. 60)

That this definition of the parameter is clearly motivated by the appeal of being able to point to the Subset Principle as the deterministic learning mechanism underlying parameter setting is clear in the following quote:

> In other words, if the values that the learning function selects on the basis of data are determined by the Subset Principle and **by nothing else** [emphasis added], then the values of a parameter must determine languages which form a strict hierarchy of subsets. That is, any two values of the parameter must determine two languages such that one is a subset of the other. (p. 61)

This seems an unnecessarily restrictive proposal as to what learning strategies might be available to the child learning a first language--a restriction placed as the ultimate consequence of assuming that the child can only use positive evidence in setting parameters.

It is interesting to note the change in the conception of a parameter represented by the work of W&M. In order to satisfy learnability demands, the parameter, rather than being associated with an array of syntactic features in a language, is associated instead with one very specific feature and is set independently, possibly for individual tokens of a lexical type, rather than for the language as a whole. This is certainly a far cry from the notion of a parameter which, once set, yields a "series of cascading effects throughout the grammar" and allows the child to acquire a cluster of related syntactic features as the result of setting just one parameter. W&M are certainly aware of this change in conception but maintain that:

> On the one hand the consequences of setting a lexical parameter would not be as broad as in the case of a language-wide parameter. On the other hand, the learnability problems might be considerably less severe. (p. 47)

It can be asked if in fact this "conflict of interests" will always be the case in formulating a parameter. That is, in the interest of maximum generality and allowing the child to learn an array of features by setting a single parameter, the formulation of the parameter will of necessity be very different from the formulation which is driven by demands of learnability and satisfying the "no-negative evidence dictum." The situation is reminiscent of the early eighties when changes in the generative linguistic theoretical framework were motivated to at least some extent by learnability considerations, and specifically by the failure of the theory in its formulation at that time, to reach explanatory adequacy (Williams, 1989).

Safir (1987) discusses this conflict between the traditional view of parameters and the lexicalized view proposed by W&M. The traditional view is motivated by the goal of avoiding undergeneralization. That is, the first language learner is assumed to learn whole clusters of syntactic properties associated with a parameter as the result of setting the value for that one parameter. On the lexicalized view, the primary motivation is to avoid overgeneralization, since it leads to a problematic learning situation whenever a subset relation exists between the strings generated by different parameter values; i.e., there is no positive evidence available to allow the learner to recover from the overgeneralization. As Safir points out, however, "...confining parameters to settings for individual lexical entries, while it alleviates the overgeneralization problem, is confronted with a potential undergeneralization problem as a result" (p. 80). This is because certain observed generalizations find no statement in the grammar.

Specifically, he points out that there may be a correlation between what a proper antecedent can be and the domain size for anaphors. Long distance anaphors must be bound by subject NPs and not by object or indirect object NPs. This generalization cannot be captured within the W&M formulation of parameters. Furthermore, with a lexicalized parameter, the fact that the proper antecedent in a particular language (e.g., Icelandic) is the same for anaphors and pronouns is not captured. That is, if anaphors must be bound by a subject, then pronouns must be free from subjects, in their respective governing categories. Thus "while the W&M approach provides a more restrictive theory of "possible parameter," the result could be an acceleration of the atomization of parameters that might render their role more descriptive" (p. 81).

This atomization can best be illustrated with an example. Recall that W&M suggest that parameter values "are associated with" lexical items in a language. It is not clear exactly what they mean by this, but if it means that each and every lexical item relevant to a particular parameter is marked for it's value of the parameter, then, for instance, every individual pronoun would have to be independently marked for its governing category and proper antecedent. That is, 'he', 'she', 'you', etc., or their equivalents in any given language, would be independently marked for the value of their governing domain and proper antecedent. The generalization that all pronouns have the same governing category and proper antecedent would not be stated in the grammar. This lack of generality is clearly at odds with the fundamental goal of any description of grammar: to capture obvious generalizations. Such a grammar fails not only in terms of explanatory adequacy, but, more fundamentally, in terms of descriptive adequacy.

A final problem of the lexicalized view of parameters is that some syntactic features which are generally presumed to be parametric ones, cannot be stated as properties of individual lexical items. The direction of complement-taking is an example. If, in any given language, complements occur to the right of verbs, we would not expect to find some verbs that exceptionally take their complements to the left. It hardly seems reasonable to encode this aspect of word order as a property of each and every verb in the language rather that stating it as a general property of the language. Such a formulation would not be a descriptively adequate one. Furthermore, there is no subset relation between different values of the parameter of direction of complement-taking, and no particular value is likely to be more or less marked than another. Safir suggests identifying those classes of parameters that do not seem to be regulated by the Subset Principle in order to improve on the W&M approach.

5. The Null Subject Parameter: Initial Settings

The null subject parameter (Rizzi, 1982), also referred to as the pro-drop parameter (Chomsky, 1981) and the AG/PRO parameter (Hyams, 1986), has been proposed to capture the variability that exists across natural languages with respect to the possibility of allowing phonetically empty subject in tensed clauses. Languages that optionally allow empty subjects, such as Italian, Irish, Spanish, ASL, and Chinese, are called null subject languages. Conversely, languages that obligatorily require overt subjects, such as English, French,and German, are said to be non-null subject languages. For the following discussion we will use the more descriptive term null subject parameter/language as it doesn't presuppose any particular formal analysis. However, in discussing particular proposals we will adhere to the terminology used by the originator of the proposal.

There have been a number of proposals for how best to formulate the null subject parameter. Each proposal provides its own answer to the questions of what features the parameter encodes, what evidence is available to the child to set it, what the initial setting is, and its relation to markedness theory. Furthermore, each of them offers an explanation of developmental stages observed in first language acquisition. We turn now to a detailed consideration of these proposals.

One proposal, first laid out in Hyams (1986, 1987a) and further discussed in Hyams (1987b) and Jaeggli and Hyams (1987) argues that children embark on the task of language acquisition assuming that the language they are learning is a null subject language. Under this view, children learning a language like Italian need not "reset" the parameter since their parametric setting "matches" the parametric setting in the adult grammar. On the other hand, children learning a language like English will necessarily have to reset the parameter to [-] null subject in order to arrive at the adult grammar.

An alternative proposal, advanced in Rizzi (1982) and supported in P. Bloom (1990, 1993), has as its premise that children begin their language acquisition with a non-null subject setting and switch to a null subject setting in case there is positive evidence in the adult grammar which contradicts the current setting. It follows from this that children learning an "English-type" language do not have to reset their parametric value, whereas children learning an "Italian-type" language do.

Finally, in a third proposal, Valian (1990a, 1990b, 1993) contends that neither a [+] null subject nor a [-] null subject value is an adequate descriptor of the child's initial state. Rather, children must have both values of the parameter available at the outset of language acquisition.

5.1. Hyams 1986: [+] Null Subject Initial Setting

It is a well attested fact that early English child language is characterized by the frequent omission of subjects, even in cases where the adult grammar necessarily requires them. The following sentences, taken from Bloom & Lahey (1978), are representative of this phenomenon.

(2)   Gonna eat supper.
        Broke this.
        Want a house like that.
        Here go again.

According to Hyams (1986, p. 26) these sentences are **"the well-formed output"** [bold added] of an early English child grammar in which the empty category in subject position is licensed because INFL is [+] pronominal. In her analysis, INFL contains two nodes: (AG)greement and (AUX)iliary. AG contains a set of features for person, number, and gender associated with the subject--generally referred to as "θ features" (following Chomsky, 1982)--and AUX minimally contains the tense specifications of the sentence. Hyams proposes that languages differ with respect to whether AG equals PRO. When this is the case (e.g. in Spanish and Italian), an empty category in subject position is allowed by virtue of PRO licensing it. Where AG is not equal to PRO (e.g. in English and French), a null subject is impossible.

Hyams proposed that a PRO analysis of AG explains the quite distinct behavior that auxiliaries and modals exhibit in English and Italian. Data from English subject-AUX inversion, VP deletion, tag questions, and verbal morphology, provide considerable evidence that the English modals and auxiliaries appear under AUX, hence constituting a separate constituent from VP. Specifically, modals are base-generated in AUX, and the auxiliaries 'be' and 'have' are raised into AUX.

Zagona (1982) and Rizzi (1982) have noted that there is no reason to believe that Spanish and Italian auxiliaries appear under the AUX constituent, since in those languages, there is no process of tag formation, negative markers cannot intercede between the auxiliary and the verb, and clitics must precede both the auxiliary and the main verb. With respect to the behavior of modals, Rizzi notes that, in contrast to English, "modals" in Italian exhibit the full range of morphological behavior that main verbs do, and thus proposes that in this language, modals are main verbs and auxiliaries appear with the main verb under VP.

Returning to the AG/PRO parameter, Hyams proposed that the impossibility of generating or raising auxiliaries into AUX in null subject languages follows from the fact that a lexical element in AUX would govern PRO, resulting in a violation of the PRO theorem, which states that PRO must be ungoverned (Chomsky, 1981).

5.1.1. The Early Grammar of Children Learning English

As previously mentioned, there is a stage during English child language acquisition which is marked by the frequent production of subjectless sentences. During this same stage, English speaking children also produce sentences with overt subjects. Thus, in addition to the utterances in (2), children also produce the following, taken from Bloom & Lahey (1978):

(3)   a. I want try that.
        b. I get money.
        c. She eat supper.
        d. My baby upstairs.
        e. Daddy sleeping.

Braine (1973) shows that children who produce subjectless sentences also produce "replacement sequences." These are versions of subjectless sentences that have been immediately expanded by adding an overt subject:

(4)   a. fall...stick fall
           b. close radio...Mommy close radio
           c. push Stevie...Betty push Stevie
           d. get...Lucy get
           e. stand up...cat stand up

According to Hyams, the fact that subjectless sentences co-occur with sentences containing overt subjects and with replacement sequences shows that the absence of lexical subjects in early English child language cannot be explained by appealing to performance limitations on sentence length. Hyams reasons that if this were the case, the child would not be capable of producing longer versions of the subjectless sentences. Moreover, Hyams claims that subjectless sentences do not correlate with complex syntactic structures--a fact that would be expected on the analysis that an increase in the cognitive load is responsible for the lack of lexical subjects. These observations have led Hyams to conclude that null subjects in early child language are entirely optional, and that this optionality provides support for a [+] AG/PRO analysis of early child language.

Returning to the English child language data, during the same period that they produce subjectless sentences, children have also been noted to systematically omit auxiliaries and modals in their speech (Bloom & Lahey, 1978):

(5)   a. I going out to playground.
           b. I not tired.
           c. I riding tank car.
           d. Where man go?
           e. I going wash my hands.

In Hyams' proposal, this lack of lexical material in AUX follows from the child's pro-drop grammar, since the presence of any lexical element in AUX would result in a violation of the PRO theorem. This is schematized below:

(6)
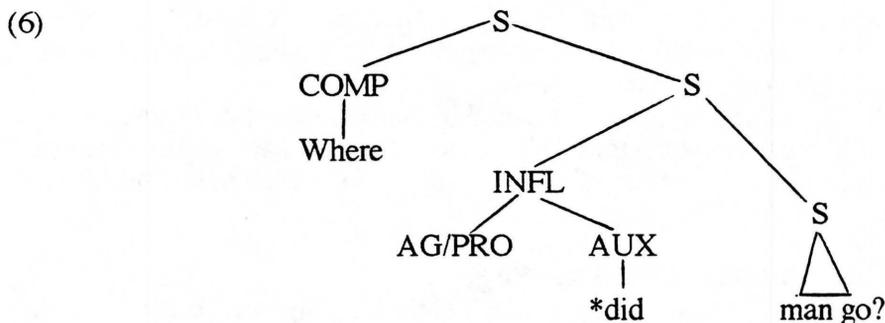
The unavailability of the AUX node in an AG/PRO grammar predicts that English speaking children will acquire modals significantly later than Italian children will, since in the second case, but not in the first case, modals are identified and analyzed as main verbs. As Hyams points out, this prediction is confirmed by the acquisition data.

## 5.1.2. The Triggering Evidence and Markedness

The claim that early English is [+] AG/PRO appears to have considerable empirical motivation. However, there remains one substantial issue to be addressed: how do children reset the parameter to [-] AG/PRO, as they clearly must in order to arrive at a correct English adult

grammar? Hyams proposes two possible triggers for resetting the parameter: (1) recognition of expletive subjects and (2) the presence of unstressed subject pronouns in the input.

By definition, expletive subjects (*it* and *there*) lack any referential and discourse function; therefore, they cannot be motivated in the same way as overt subject pronouns in [+] AG/PRO languages. They must, then, be grammatically motivated and indicative of the fact that English subjects cannot be null. Hyams hypothesizes that when children notice the presence of expletives in their input, a parametric change from [+] AG/PRO to [-] AG/PRO is triggered.

The second potential trigger is the presence of unstressed subject pronouns in the input. Hyams supposes that circumstances (e.g., pragmatic factors such as emphasis, change of topic, etc.) which require an overt subject pronoun in Italian would require a stressed pronoun in English. Hence, if children notice that in their input pronouns are used even when pragmatic factors don't require them, this would indicate to the child that subjects are necessary.

The grammatical strings generated by the [+] and [-] setting of the proposed AG/PRO parameter stand in a set-theoretic relation to each other traditionally referred to as "an overlapping" relation. The two "languages" have a non-empty intersection which does not extend to one being included in the other. The relationship can be depicted as:

(7)
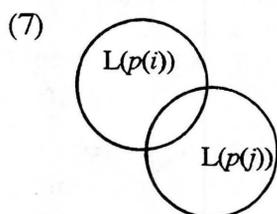
Let $p(i)$ designate a [+] setting of the AG/PRO parameter and $p(j)$ a [-] AG/PRO parameter setting. The languages (strings of grammatical sentences) generated by each of these settings can then be represented as $L(p(i))$ and $L(p(j))$. Each parameter setting generates strings shared by the two settings and, in addition, strings allowed in the language designated by one parameter setting but not the other. Syntactically null subjects will be compatible with the (i) setting but not the (j) setting. Overt expletives will be compatible with the (j) setting but not the (i) setting. Overt referential subjects will be compatible with both settings.

5.1.3. Some Comments on Hyams 1986

There were a number of empirical problems in Hyams analysis which led her to a reanalysis of the parameter which we will consider in detail below. However, one problem was not identified, and since it is relevant to our discussion of parameter setting, we will consider it now. The problem concerns the proposal that overt expletives in non-null subject languages are the trigger for resetting the value of the parameter from the incorrect initial assumed [+] setting to the correct [-] setting when the child is actually learning a non-null subject language. What this requires, cross-linguistically, is that 1) all null subject languages require null expletives, and 2) all non-null subject languages require overt expletives.

(1) is necessary so that the occurrence of overt expletives in the data can serve, universally, as positive evidence for the child that the target language is a non-null subject language. If a null subject language allowed the possibility of either an overt or null expletive, then the existence of expletives cannot serve as positive evidence that the target language is a non-null subject one. Overt expletives are compatible with both settings of the parameter. In this situation, only the absence of null expletives in the data differentiates the null subject from the non-null subject language, and this kind of indirect negative evidence, by assumption, is not evidence that the child can use. In other words, if languages which allow null referential subjects do not *require* null expletives but only *allow* them, then we have a subset relation between the two languages with the superset language represented by the [+] AG/PRO setting of the parameter. As we have seen, in this situation, if the child's initial setting is the one which generates the superset language, then there is overgeneralization from which the child cannot recover on the basis of positive evidence only.

Whether or not all null subject languages require null expletives is an empirical question which can be resolved by looking at the relevant cross-linguistic data. To our knowledge this has not been done. It is nonetheless an implicit empirical prediction of Hyams' proposal which bears further investigation.

The second implicit requirement of Hyams' proposal, that all non-null subject languages require overt expletives, follows from the fact that without this crucial trigger, a child learning a non-null subject language would not be able to recover from the initial incorrect setting of the parameter. This implicit requirement does face a strong empirical challenge, as there are non-null subject languages (German) which do not require overt expletives. A child learning German, under Hyams' proposal, would never get the necessary data to trigger the appropriate adult setting of the parameter.

The implicit predictions pointed out above are indicative of a general problem in the parameter setting model of language acquisition. If the universal nature of the language acquisition device is to be maintained, the relevant data for setting and resetting parameters must be universally applicable and thus must be informed by a comprehensive study of precisely what co-occurrences there are cross-linguistically with regard to the relevant features of the parameter. Problems in Hyams' analysis stem from the fact that the two languages (English and Italian) on which the "universal" analysis was built are a non-null subject language that requires overt expletives and a null subject language that requires null expletives. The situation is likely to be more complex when a variety of languages are considered. Fortunately, recent reanalyses of the parameter (Jaeggli & Safir, 1989) are based on just such a broad consideration of actual cross-linguistic variation.

In very general terms then we can see that if a subset relation is generated by two values of a parameter involving a particular syntactic feature, then, if another feature is added to the same parameter, there will be restrictions on possible cross-linguistic co-occurrences of the two features if the subset relation is to be maintained universally. As we saw, similar considerations led Wexler and Manzini to propose that in fact the features encoded by a parameter needed to be stated in very limited terms in order to satisfy learnability criteria--specifically the Subset Condition. As the number of features associated with a particular parameter increases, the possibility of the languages generated by different values of the parameter standing in an overlapping relation to one another also increases. And the precise way in which the languages overlap way vary from one language to another, making it difficult to identify what will serve as the "universal" triggering data for parameter setting. Without the identification of such a universal trigger, one has to resort to proposing that children learning different languages engage in a kind of inductive reasoning in setting parameters. The parametric model, however, was motivated by the desire to avoid having to postulate these kinds of general inductive learning strategies.

5.1.4. Accessibility of a [+] Null Subject Setting

Before turning to a consideration of Hyams' reanalysis of the null subject parameter, let us consider the question of markedness under the 1986 analysis. As we have seen, Hyams proposed that the languages generated by the values of the AG/PRO parameter stand in an "overlapping" relation and, thus, the Subset Condition is not met and the Subset Principle cannot apply. On the question of which setting of the parameter is more marked Hyams remains uncommitted. There is one approach to markedness which defines markedness in terms of the length of successive child grammars leading to the final adult grammar. In this approach, the child's initial assumption, whatever it may be, is by definition the least marked. This definition is a very circular one, and Hyams does not appeal to it. She prefers to "leave open the question of whether different parameter options are ranked with respect to markedness, particularly in regard to the AG/PRO parameter" (Hyams 1986, p. 162).

However, some explanation is still required to explain the fact that the [+] setting of the parameter is more "accessible" to first language learners as evidenced by the fact (according to Hyams) that they initially assume that setting. The explanation which Hyams offered is what she called the *Isomorphism Principle* which she formulated as follows:

All else being equal, the least complex grammatical system is the one which allows for the greatest degree of isomorphism between the various levels of representation, D-structure, S-structure, PF, and LF. (p. 162)

Briefly, a grammar involving a difference between any two levels in the representation of a sentence is less preferred to one without such a difference. By assumption, all subjects are represented at D-structure as sets of agreement features without a phonological matrix, so the [-] null subject language, which requires such a matrix at S-structure, is less isomorphic than the [+] null subject language in which the subject can remain at S-structure, as at D-structure, without a phonological matrix.

As we have seen, Hyams' proposal is empirically motivated and offers an account of some aspects of the early stages of English grammar (see Hyams, 1986 for additional acquisition data from German and Italian which is consistent with her view). As it stands, however, her proposal suffers from several empirical inadequacies, which are discussed in Hyams, 1987a. For one, Hyams predicts that children acquire modals at roughly the same stage at which they abandon null subjects and finds confirmation of this hypothesis in child acquisition data.[3] However, as Hyams herself points out, the infinitival marker 'to' also emerges alongside the modals, a fact which cannot be explained within the framework outlined above. Also, in her analysis, Hyams argued that null subjects are possible in languages which have, in some sense, rich agreement features (often realized in the form of rich overt inflection). Children acquiring richly inflected languages like Italian do acquire the inflectional system at a very early age. However, English speaking children use null subjects despite the fact that English verbal morphology is not yet acquired. This developmental difference also remains unexplained.

These inadequacies led Hyams to propose a reanalysis of the null subject phenomenon based on parametric variation with regard to a property of language called Morphological Uniformity. Let us now look at this more recent proposal in some detail.

5.1.5. Hyams: A Reanalysis of [+] Null Subject

A basic component in Hyams (1987a) is the distinction between licensing conditions for empty categories and identification conditions for these empty categories (discussed in Jaeggli & Safir, 1989). The idea is the following: null subject languages such as Chinese and Italian share the property that their morphological system is uniform. That is, whereas Chinese has only underived inflectional forms, Italian possesses only derived ones. This is exemplified in Atkinson (1992), and repeated in (8) below:

(8)	Chinese		Italian	
(wo)	gongzuo	(io)	lavoro	'I work'
(ni)	gongzuo	(tu)	lavori	'You work'
(ta)	gongzuo	(lui/lei)	lavora	'He/she/it works'
(women)	gongzuo	(noi)	lavoriamo	'We work'
(mimen)	gongzuo	(voi)	lavorate	'You (pl.) work'
(tamen)	gongzuo	(loro)	lavorano	'They work'

English, on the other hand, exhibits a mixed paradigm by virtue of only some of its forms being inflected:

(9)	I	sleep
	You	sleep
	He/she/it	sleeps
	We	sleep
	You	sleep
	They	sleep

Jaeggli and Sapir (1989) proposed that the correlation between uniformly inflected (or uninflected) paradigms and the appearance of null subjects on the one hand and mixed paradigms and the absence of null subjects on the other indicates that the generality is that null subjects are allowed in languages that are morphologically uniform. They offer the following formulation of the null subject parameter, which determines when null subjects in tensed sentences are licensed:

> Null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms. (p. 29)

It should be emphasized, however, that not all languages that are morphologically uniform allow null subjects, as is the case in German:[4]

(10)

*(∅)/Ich	arbeite	'I work'
*(∅)/Du	arbeitest	'You work'
*(∅)/Er	arbeiret	'He/she it works'
*(∅)/Wir	arbeiten	'We work'
*(∅)/Ihr	arbeiten	'You (pl.) work'
*(∅)/Sie	arbeiten	'They work'

Thus, although morphological uniformity is a necessary condition for allowing null subjects, it is not a sufficient one. Morphological uniformity only establishes the possibility of null subjects. For referential null subjects to appear in a language, Jaeggli and Sapir propose that null subjects must also be identified. To exemplify this, let's return to the Italian paradigm in (8) above. In Italian, null subjects are licensed by virtue of its having a morphologically uniform inflectional paradigm. It was also noted that Italian is a language with rich inflectional morphology. According to Hyams, it is the features in INFL (or AGR) that eventually get "spelled out" as verbal inflections which permit identification. In sum, then, null subjects occur in Italian because: (1) they are licensed via morphological uniformity, and (2) they are identifiable by means of rich agreement.

Let us now examine the case of Chinese. Chinese is a paradigmatic case of a language without verbal inflections. In addition, Chinese is broadly characterized as a member of a set of discourse-oriented languages--in contrast to Italian, which is said to belong to the set of sentence-oriented languages. Huang (1982) maintains that the notion of "discourse-topic" interacts with purely syntactic processes to identify empty subject positions. This is schematized in (10) below:

(10) $[_{Top} e_i] [_S e_i$ gongzuo]

According to Huang, "Top" is an operator in an A-bar position which binds--and hence identifies-- e_i, the variable in the subject position. Drawing on this observation, Hyams proposes that there are two mechanisms whereby null subjects are identified: "agreement" identifies null subjects in sentence-oriented languages, and "topic" identifies null subjects in discourse-oriented languages.

In one respect the morphological uniformity approach to the presence or absence of null subjects is a distinct improvement over previous analyses, in that it is based on data from a fairly wide range of languages including Japanese, Chinese, Spanish, Italian, German, Irish, Danish, Swedish, French, Arabic, & ASL. As noted earlier, this would seem to be a minimal requirement for a hypothesized "universal" explanation of any syntactic feature. On the other hand, the analysis does lack intuitive appeal. Why should it be the case that morphological uniformity is related to the presence or absence of overt subject pronouns? There seems to be no logical connection between the two. Under the previous analysis, null subjects were tied to rich inflection on the verb, which makes sense as this inflection was what identified the referent of the null subject. No such logical association exists in the reanalysis offered by Jaeggli and Safir. Perhaps a future analysis will offer a formulation that is both consistent with the cross-linguistic data and has greater logical appeal.

5.1.5a. Hyams' Reanalysis of Early Child Language

Hyams notes that children learning German, English, or Italian omit subjects even in those cases where the adult language obligatorily requires them. She concludes from this that children must be treating the language as if it were morphologically uniform, and proposes the parameter responsible for the null subject to be as follows:

(12) The Morphological Uniformity Parameter:
        a. [+] uniform
        b. [-] uniform

From this, an obvious prediction follows: the appearance of verbal inflection in English speaking children should correlate with the disappearance of null subjects. This is because children's awareness that English has a mixed inflectional paradigm would cause them to realize that null subjects are not licensed, and hence cannot occur in the language. Hyams (1987a, p. 3) maintains that this prediction is borne out, pointing to Guilfoyle (1984) who "has noted that English speaking children begin productive use of verbal inflection at around the time they shift from a null subject to a non-null subject grammar."

However, there remains a problem in the analysis. If agreement inflections are not present for English-speaking children, during their null subject stage, what identifies the null subjects? Hyams' way out of this tangle is to claim that English speaking children possess, in the relevant aspects, "Chinese grammars." Recall that Hyams proposes that null-subjects in Chinese are identified via "topic." Hyams extends this notion to early English grammars to say that null subjects are identified using the same "topic identification" mechanism.[5]

5.1.5b. Markedness under Morphological Uniformity

Before turning to the next proposal, we will consider Hyams reanalysis in terms of the question of markedness. Hyams maintains, based on developmental data, that children's initial setting for the null-subject parameter is [+]. That is, they assume that the target language is morphologically uniform. Under the reanalysis, the two values of the parameter now generate "languages" which stand in a subset relation to each other. A morphologically uniform language will have morphological forms that are the "same" in that they are all either 1) uninflected (e.g., Chinese) or 2) inflected (e.g., Italian). The morphologically non-uniform language will have some forms that are uninflected as well as some that are inflected. Hyams' reanalysis then involves a parameter which satisfies the Subset Condition so the Subset Principle can apply, which establishes the [+] morphologically uniform setting as the unmarked one, as it is the more restrictive hypothesis. In the event that the child hypothesizes a morphologically uniform language, there will be forms in the data (inflected or uninflected) which will serve as positive evidence to the child that the target language is not morphologically uniform in the event that it isn't. If the child were to hypothesize the [-] morphologically uniform setting there would never be disconfirming evidence if that setting were incorrect. Although all the forms presented would in fact be morphologically uniform--either all inflected or all uninflected--the child could not know that there wasn't non-uniformity in the paradigm without some strategy of deciding at some point that the data presentation was complete. This is equivalent to using indirect negative evidence as the child is basically concluding that something doesn't exist in the language on the basis of not having heard it up to some point in the acquisition process. So, from a learnability-theoretic perspective, with only positive evidence available to set the parameter, the child must, and according to Hyams' analysis of the data does, assume [+] morphologically uniform as the initial setting for the null subject parameter.

Thus far, we have outlined in considerable detail the proposals Hyams has put forth to account for the occurrence of missing subjects in early child grammars. There is however an alternative explanation for such phenomenon, one that regards the omission of subjects not as being linguistically motivated but rather as a consequence of processing constraints. Let us now consider this proposal in detail.

5.2. Bloom: [-] Null Subject Initial Setting

In section 5.1.1, it was pointed out that Hyams rejects a processing account of null subjects, claiming that the occurrence of subjects is independent of utterance complexity. Recently, this view has been challenged by P. Bloom (1990, 1993), who performed a series of analyses on the data of three children studied by Brown (1973). Bloom demonstrates convincingly that his results are consistent with a processing explanation. He takes these findings as evidence for the claim that children embark on the task of language acquisition with a [-] null subject setting.

Before turning to Bloom's analysis, it is important to understand the notion of "performance limitations" assumed in the 1990 study. He explains it as follows:

> For very young children, there is an imperfect mapping from what they intend to say to what they actually say. So, constituents that are present at some deeper level of representation occasionally do not appear in children's utterances. (pp. 491-492)

Thus, Bloom begins with the premise that children have the same grammatical capacities that adults do, and that their production is not a reflection of a "deficient" competence, but rather the result of performance limitations.

Bloom (1990) finds abundant empirical evidence from previous studies supporting the processing account of null subjects. He cites Brown and Fraser (1963) and Ervin (1964) who show that the length of young children's utterances in an imitation task is predicted by the length of the child's spontaneous utterances, and not by length of the adult sentence which they are asked to imitate. Bloom interprets this result as indicating that children's sentences are frequently short not because of inadequate grammars, but because they have a general difficulty in producing long sentences.

As additional evidence, Bloom reports on a study conducted by L. Bloom (1970), in which she examined all the utterances of a 22-month-old child that contained the verb 'make'. Her prediction--assuming a processing limitation account--was that if a subject was present in the utterance, some other element would fail to surface. Analysis of the child's utterances confirmed her hypothesis: when the utterance contained an overt subject, either the verb, the object, or the adverbial phrase failed to surface. Thus, young children omit all constituents, not just subjects--a fact nicely accounted for in terms of general performance limitations.

A third study cited by Bloom is that of Mazuka et al. (1986), where it is attested that some children go through a stage in which they neither include nor omit subjects from sentences. Instead, children phonetically reduce subjects to a schwa. Again, this outcome is expected if children have processing limitations, but not if children have null subject grammars. The reason for this is that children, knowing that their grammars require overt subjects, attempt to produce them but partially fail to do so.

Although the evidence presented above favors a processing account for subject omission, Hyams (1987b) brings up two interesting observations which need to be accounted for. First, Hyams maintains that the absence of subjects in children cannot be due to processing difficulties since "lexical subjects may be absent in simple utterances as they may in more complex sentences" (p. 10). Second, there appears to be a striking asymmetry with respect to subject and object omission. That is, while missing subjects are abundant in child language, missing objects are remarkably scarce--a phenomenon that, in Hyams' view, cannot be explained via non-grammatical accounts. Bloom (1990, 1993) addresses these issues.

5.2.1. The Relationship between Subject Omission and VP Length

To address Hyams' first concern above, Bloom (1990) compared the average length of the VP's produced by children in sentences containing overt subjects with those containing null subjects. The logic behind this is that if children's subjectless sentences are the result of processing difficulties, then they should omit subjects more often from longer than from shorter utterances.

The subjects for this study were three children (Adam, Eve, and Sarah) studied by Brown (1973), whose language samples are transcribed and stored in the CHILDES data base (MacWhinney & Snow, 1985). The structures that Bloom concentrated on were of two types: (i)

sentences with past tense verbs, and (ii) sentences with verbs that denote cognitive states or involuntary acts (e.g., need).[6] Questions, statements with 'no' or 'don't', statements where the verb was part of an embedded clause, rote imitations of adult speech, imperatives, and some questions with null subjects which were acceptable in adult speech (e.g., want a cookie?) were excluded from the analysis.

The length of the VP was calculated by counting the number of words from the verb onward, until it was clear that they were not part of the VP. Thus, Bloom points out that an utterance such as 'I goed to the bathroom, Mommy' was counted as having a VP with three words, and not four.

Bloom compared the mean length of VPs in sentences with and without subjects for all three children and concluded that in all but one,[7] the difference between the two means, as measured with a one-tailed t-test, was statistically significant. This clearly points to the fact that null subjects are the result of processing constraints (see Bloom, 1990 for an alternative explanation of these results which favors a null subject analysis and its subsequent refutal based on a contrast between subject pronouns and non-subject pronouns).

5.2.2. The Subject/Object Asymmetry in Processing

Recall that Hyams' second objection to a processing account for null subjects is the observation that subjects tend to be more frequently omitted than objects. Bloom (1990) reasons that if the subject/object asymmetry is due to differences in processing limitations--that subject position is more vulnerable to processing limitations--we would expect to see other differences between subjects and objects. First, since use of pronouns can reduce processing load compared to use of nonpronouns (because the former are shorter), the proportion of pronoun NPs in subject position should be greater than the proportion of pronoun NPs in object position. Second, we would expect overt nonpronoun subjects to be shorter than overt nonpronoun objects. As Bloom notes, both of these predictions are confirmed (see Bloom, 1990 for a discussion of the results).

In addition to these findings, Bloom (1993) cites several studies which show that the subject/object asymmetry can be accounted for without resorting to a grammatical explanation. The first one comes from Bever (1970), where he notes that the position of a "heavy constituent" in a sentence affects its acceptability. Thus, 'John walked briskly in a slightly northerly direction' is preferred over 'John walked in a slightly northerly direction briskly'. This suggests that there is a bias to place "heavy constituents" near the end of the sentence. Bloom suggests that this same bias might favor objects over subjects in children's production.

A second argument comes from a study by Gerken (1991), where she reports that 2-year-old children tended to omit articles more frequently from subject NPs than from object NPs (31% vs. 18%) when engaged in a sentence imitation task. This same result was obtained by examining children's spontaneous utterances, where children omitted articles 31% of the time from subject NPs as opposed to 14% from object NPs. This result, Bloom argues, is consistent with a processing account of the subject/object asymmetry.

Finally, Bloom presents an additional argument favoring a non-syntactic explanation for the subject/object asymmetry. He suggests that the frequent subject omission could be due to pragmatic factors. It has been traditionally known that whereas subjects convey "known" or "old" information, objects typically convey "new" information (Leech & Svartvik, 1975). Bloom (1990) reasons that if children's production is indeed constrained by processing limitations, "the best way to cope with processing limitations while communicating successfully is to omit or reduce subjects, not objects " (p. 501). Since redundant information is more often expressed in subjects than in objects, subjects will be more frequently omitted than objects.

To conclude, then, Bloom, following Rizzi (1982), takes the view that children initially assume a [-] value for the null subject parameter and presents abundant evidence to support the claim that the omission of null subjects in early child grammar can be readily explained as the result of performance limitations.

In terms of the question of markedness, the [-] null subject value would be the unmarked value, and this is the value initially furnished by UG. The relationship holding between a [+] null subject and a [-] null subject language is a subset one, so the Subset Condition is met and the

Subset Principle applies to provide the child with the initial, [-] null subject parameter. This parameter setting is compatible with the input data, and it is the smallest among the languages compatible with such data.

Thus far, we have discussed two possible initial settings for the null subject parameter: one in which the initial setting is [+] null subject (or [+] morphologically uniform to be precise) and one in which the initial setting is [-] null subject. One might suppose that given a parameter with only two possible values, there would only be two proposals for what the initial setting must be. There remains, however, a third possibility which is that the child embarks on the acquisition process with both values of the parameter available. This radically different view was proposed by Valian (1990a, 1990b).

5.3. Valian: Both Parameter Values Available

Valian (1990a, 1990b) proposes that neither a [-] null subject nor a [+] null subject value can adequately characterize children's initial assumption about the setting of the null subject parameter. Her argument is motivated by two distinct observations: (a) parser limitations, and (b) the existence of misleading input. We will begin by summarizing the arguments Valian presents against a [-] null subject value as the child's initial state.

5.3.1. The [-] Null Subject Value

Valian points out that starting the child off with a [-] null subject value is advantageous from the standpoint of learnability. It is advantageous because we guarantee that all that the child needs to reset the parameter is positive evidence in the input which contradicts the current setting. However, Valian maintains that this outcome is not possible. Her arguments rest on the notions of parser limitations and misleading input.

5.3.1a. The Parser Failure Argument

Suppose the child starts by assuming that the target is a [-] null subject language (as in English). Suppose further that the child is being exposed to a [+] null subject language (e.g., Italian). Valian argues that incoming subjectless sentences will be uninterpretable by the child because her parser is fed by a grammar in which sentences are labeled as such if and only if they have overt subjects. Valian (1990a) explains the role of the parser in the acquisition of the correct parametric value by a Spanish child as follows:

> If the child's parser is solely fed by an English grammar, she will reject the null subject string as nongrammatical, and never switch to the Spanish value. Her parser will not assign the label "sentence" to anything without a surface subject. (p. 123)

Furthermore, Valian argues that we cannot simply assume that the parser automatically labels all strings without subjects as "sentences," as this would create a problem for children learning English with a correct [-] null subject setting. They would be mistakenly led to change that setting on the basis of misleading input (see below). According to Valian, then, positive evidence is not sufficient to trigger a parametric change since children's [-] null subject grammar would not allow them to appreciate the significance of the input received.[8]

5.3.1b. The Misleading Input Problem

This constitutes Valian's most articulate and best developed argument. The concept of misleading input is quite simple. Recall from our previous discussion that children need to be exposed to certain types of triggering evidence to set parameters correctly. So, a Spanish child starting off with a [-] null subject grammar would take sentences with null referential subjects as evidence for the need to reset the parameter to [+] null subject. There are cases, however, in which the type of input the child gets mimics the trigger in some respect, as exemplified below:

(13)   Wanna cookie?
         Bet you can't stack those blocks.
         Gonna get your lunch, and I'll be right back.

These cases are labelled "misleading input" because they can potentially cause children hearing English to arrive at the wrong conclusion about their target language. Their grammar dictates that sentences are required to have overt subjects. Hearing repeated instances of sentences as in (13), without overt subjects, could cause them to incorrectly conclude that they are learning a [+] null subject language and result in an unnecessary switch in the parameter.[9]
         In view of these facts, Valian proposes that children's acquisition of the null subject parameter cannot be guaranteed by proposing a [-] null subject value. Let us now turn to the alternative of starting the child off with a [+] null subject grammar.

5.3.2. The [+] Null Subject Value

Recall that Hyams (1986), based on the empirical fact that null referential subjects are common in children's production, claims that all early grammars are initially [+] null subject. She suggests that children's initial value is reset on the basis of the presence of expletives in the language. Valian (1990b) claims that Hyams cannot adequately describe children's initial grammar because her resetting account crucially depends on the assumption that children hear fully grammatical input. Let us now look at Valian's argument more closely.

5.3.2a. Misleading Input and Non-Parser Failure

Take the case of children with a [+] null subject parameter setting. If their target language is a [+] null subject language, they will not need to reset the parametric value. If, on the other hand, the child is learning English, the parameter will have to be reset to [-] null subject in order to arrive at the adult grammar. In this latter case, given that the child's parsing mechanism is fed with a grammar that labels subjectless utterances as sentences of the language, hearing null subject sentences will cause the child to incorrectly retain the hypothesis that the target language is [+] null subject. We might, at this point, be tempted to think that expletives will provide the child with the relevant evidence to reset the parameter. Valian maintains that this is not a possibility because the child's parser will supply her with a referential interpretation for the expletive pronoun. This is because the expletive interpretation is only available as an entrained consequence of the [-] null-subject setting of the parameter.

5.3.3. Dual Value Solution and Hypothesis-Testing Mechanism

Thus far, we have considered two developmental accounts of the null subject phenomenon. Both of these have children start with one value of the parameter, and subsequently require the child to reset it just in case the initial parametric value does not match the adult state. Valian observes that parameter resetting under either account is indeed an impossible task and proposes a dual-value setting of the null subject parameter. She argues that children must have both the [-] and the [+] parameter values available.
         An advantage of the dual-value solution, Valian argues, is that it solves the parser failure problems discussed above. Let us consider how this would work. Recall from our discussion that children starting off with a [-] null subject value will fail to assign the correct interpretation to an incoming Italian subjectless utterance because the parser reserves the label "sentence" for those strings that have overt subjects. Valian maintains that this problem is overcome with a dual-setting value because now the child's parser, being fed with a grammar that also allows subjectless sentences, will correctly label the null subject utterances as fully grammatical sentences. A similar case can be made for the interpretation of "expletives." In Valian's view, an initial [+] null subject value will not allow the child to interpret expletives, so a referential interpretation will be forced. However, a parser fed with a [-] null subject grammar in which expletive subjects are interpretable as semantically null will overcome this problem.
         Valian notes that although the dual-value solution solves the parsing problem, it raises another problem: an English child would never be able to fix the parameter to [-] null subject

because there is no input that would ever count as evidence that the null subject parameter is incorrect. This is because every sentence that is interpretable by the non-null subject grammar is also interpretable by the null subject grammar. Valian's solution to this problem is to invoke a hypothesis-testing mechanism that decides which of the values of the parameter best fits the input. Now let us examine how such a mechanism works.

As stated above, both values of the parameter are available to children. Hence, when they hear a subjectless sentence, they have to determine if it is a fully grammatical sentence of the language or if it is just "acceptable," able to be produced and comprehended for reasons that fall outside syntax. Similarly, when they hear expletives, they must decide whether they are true expletives or referential subjects. To make such a decision, the child has to evaluate and weigh evidence. For this process to be carried out, Valian invokes a psychological mechanism that will search for clues that will allow the child to decide between a null subject language and a non-null subject language. One way the child could do this is by comparing adult usage of subjects in different positions in a sentence. If the child is faced with a prototypical null subject language, subject omissions will appear across all types of clauses: matrix and embedded. The child will hence conclude that the correct parametric value is [+] null subject. For the child learning a [-] null subject language, the situation is similar. Thus, the child presented with an English-type language will notice that the distribution of null subjects is skewed since they occur in matrix clauses but not in embedded clauses. This will lead the child to conclude that null subject sentences in English are only acceptable, and this will lead to a [-] setting of the null subject parameter.

6. Conclusion

The notion of a parameter has undergone some radical revision since its initial formulation within Principles and Parameters theory. As we have seen, these revisions have been motivated by the learnability considerations which are central to modern theoretical linguistics. Ideally, a parameter would be formulated as a cluster of syntactic features that can be learned simultaneously, and universally, once the relevant, simple trigger in the data is encountered by the first language learner. This ideal view is consonant with the dominant view within Principles and Parameters theory that language learning mechanisms are deterministic and automatic. This ideal, however, seems to be at odds with basic learnability requirements, and in some cases, with the empirical facts of cross-linguistic variation. As we have seen, learnability requirements and cross-linguistic variation often call for a very different formulation of parameters than is called for by considerations of maximum generalizability in describing any particular language.

Nonetheless, it is a sign of robustness that a theoretical paradigm contains within it the impetus for its own revision. In this case the criterion of explanatory adequacy serves that function. Chomsky (Gleidman, 1985) has said:

> Undoubtedly the principles of universal grammar that we currently theorize are wrong. It would be a miracle if we were right this early along. But the principles are of the right type, and we can now begin to test our present system with complex examples to see what is wrong and to make changes that will improve our theory...." (p. 373)

And one can hope that, as long as the fundamental goals of inquiry are valid and assumptions are revised in light of empirical investigation, the process of theory building will ultimately guide the investigation towards identification of those "principles of universal grammar" as well as the ways in which individual languages vary with regard to those principles.

THE AUTHORS

Paola Dussias is a doctoral candidate in the SLAT program, majoring in L2 Analyis and minoring in L2 Processes and Learning. Her interests include language attrition, code switching and childhood second language acquisition.

Martha Schulte-Nafeh is also a doctoral candidate in the SLAT program with a specialization in L2 Analysis. She has a minor in L2 Use with a focus on sociolinguistic issues relevant to the Arabic-speaking world.

Program in Second Language Acquisition and Teaching
Modern Languages Building, Room 445
University of Arizona
Tucson, AZ 85721

e-mail addresses: pdussias@ccit.arizona.edu
                              msnafeh@ccit.arizona.edu

NOTES

[1]In more recent research, it has been claimed that negative evidence in the form of "noisy feedback" from caregivers is available to children in first language acquisition. Interested readers are referred to Marcus (1993) for a review and refutation.

[2]Oerhle (1985) offers a formal explication of how implicit negative evidence might play a role in the acquisition of gaps in morphological paradigms as well as partially regular syntactic alternations.

[3]Valian (1990a) presents a detailed study investigating the appearance of modals in the speech of a group of English-speaking children, where she concludes that it is not the case that a surge in modal-usage accompanies the disappearance of null subjects.

[4]Atkinson (1992) points out that this characterization is indeed an oversimplification of German, which actually does allow null subjects and null objects in topic position.

[5]See Atkinson (1992) for a discussion of problems raised by this proposal.

[6]Bloom calls these "nonimperative" verbs as they almost never appear in the imperative.

[7]The exception was Sarah, in the past verb condition. However, when both verb types were counted together, Bloom found a significant difference, suggesting that the non-significance in the past verb condition could be the result of the small sample size.

[8]See Kim (1993) for a solution to this problem.

[9]See Roeper and Weissenborn (1990) for a proposal of subordinate clauses as trigger domains.

REFERENCES

Angluin, D. (1978). On the complexity of minimum inference of regular sets. *Information and Control, 39,* 337-350.

Angluin, D. (1980). Inductive inferences of formal languages from positive data. *Information and Control, 45,* 117-135.

Atkinson, M. (1992). *Children's syntax.* Oxford: Blackwell.

Berwick, R.C. (1985). *The acquisition of syntactic knowledge.* Cambridge, MA: MIT Press.

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

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

Bloom, L., & Lahey, M. (1978). *Language development and language disorders.* New York: John Wiley & Sons.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry, 21,* 491-504.

Bloom, P. (in press). Grammatical continuity in language development: The case of subjectless sentences. *Linguistic Inquiry.*

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.

Brown, R., and Fraser, C. (1963). The acquisition of syntax. In C.N. Cofer & B. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes.* New York:McGraw-Hill.

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. Hayes (Ed.), *Cognition and the development of language* (pp. 11-54). New York: Wiley.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1982). Some concepts and consequences of the theory of goverment and binding, *Linguistic Inquiry Monograph, 6.* Cambridge, MA: MIT Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use.* New York: Praeger.

Ervin, S. (1964). Imitation and structural change in children's language. In E. Lenneberg (Ed.), *New directions in the study of language.* Cambridge, MA: MIT Press.

Gerken, L.A. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language, 30,* 1-21.

Gliedman, J. (1983). An interview with Noam Chomsky. In V. Clark, P.A. Eschholz & A.F. Rosa (Eds.), *Language: introductory readings* (4th ed.)(pp. 366-375). New York: St. Martin's Press.

Gold, E.M. (1967). Language Identification in the limit. *Informatin and Control, 10,* 447-474.

Guilfoyle, E. (1984) The acquisition of tense and the emergence of lexical subjects. *McGill Working Papers in Linguistics, 2,* 20-30.

Huang, C-T.J. (1982). *Logical relations in Chinese and the theory of grammar.* Unpublished doctoral dissertation, MIT.

Hyams, N. (1986). *Language acquisition and the theory of parameters.* Dordrecht: D. Reidel.

Hyams, N. (1987a). The setting of the null subject parameter: A reanalysis. Paper presented at the Boston University Conference on Child Language Development, Boston, MA.

Hyams, N. (1987b). The theory of parameters and syntactic development. In T. Roeper & E. Williams (Eds.), *Parameter setting.* Dordrecht: D. Reidel

Jaeggli, O., & Hyams, N. (1987). Morphological uniformity and the setting of the null subject parameter. *NELS, 18.*

Jaeggli, O., & Safir, K. (1989). The null subject parameter and parametric theory. In O. Jaeggli & K. Safir (Eds.), *The Null Subject Parameter.* Dordrecht: Kluwer.

Kim, J.J. (in press). Null subjects: Comments on Valian (1990). *Cognition.*

Lasnik, H. (1989). On certain substitutes for negative data. In R. Matthews & W. Demopoulos (Eds.), *Learnability and linguistic theory* (pp. 89-105). Dordrecht: Kluwer/Academic Press.

Leech, G., & Svartvik, J. (1975). *A communicative grammar of English.* Singapore: Longman Singapore.

MacWhinney, B., & Snow, C. (1985). The Child Language Data System *Journal of Child Language, 12,* 271-296.

Marcus, G.F. (1993). Negative evidence in language acquisition. *Cognition, 46,* 53-85.

Matthews, R.J., & Demopoulous, W. (Eds.) (1989). *Learnability and linguistic theory.* Dordrecht: Kluwer/Academic Press.

Mazuka, R., Lust, B., Wakayama, T., & Snyder, W. (1986). Distinguishing effects of parameters in early syntax acquisition: A cross-linguistic study of Japanese and English. *Papers and Reports on Child Language Development, 25,* 73-82.

Oehrle. R.T. (1985). Implicit Negative Evidence. Unpublished manuscript. Unpublished manuscript, Department of Linguistics, University of Arizona, Tucson.

Pinker, S. (1990). Language acquisition. In D.N. Osherson & H. Lasnik (Eds.) *An invitation to cognitive science: Vol. 1. Language.* Cambridge, MA: MIT Press.

Rizzi, L. (1982). *Issues in Italian Syntax.* Dordrecht: Foris.

Roeper, T., & Weissenborn, J. (1990). How to make parameters work: Comments on Valian. In L. Frazier & J. de Villiers (Eds.), *Language Processing and Language Acquisition.* Dordrecht: Kluwer Academic.

Roeper, T., & Williams, E. (Eds.) (1987). *Parameter Setting.* Dordrecht: Reidel.

Safir, K. (1987). Comments on Wexler and Manzini. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (pp. 41-76). Dordrecht: Reidel.

Valian, V. (1990a). Logical and psychological constraints on the acquisition of syntax. In L. Frazier and J. De Villiers (Eds.), *Language Processing and Language Acquisition.* Dordrecht: Kluwer Academic.

Valian, V. (1990b). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition, 35,* 105-122.

Wexler, K., & Manzini, M.R. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (pp. 41-76). Dordrecht: Reidel.

Williams, E. (1987). Introduction to T. Roeper & E. Williams (Eds.), *Parameter Setting* . Dordrecht: Reidel.

Zagona, K. (1982). Government and proper government of verbal projections. Unpublished doctoral dissertation, University of Seattle, Washingtion.