Nina Hyams
Kenneth Wexler

On the Grammatical Basis of Null Subjects in Child Language

Subject omission is a pervasive and extensively studied property of early child language. Thus, young children frequently produce sentences such as those in (1) (from Bloom, Lightbown, and Hood 1975, Bloom 1970).

(1) Shake hands.
Turn light off.
Want go get it.
Show Mommy that.
Not making muffins.

In an attempt to explain the missing subject phenomenon, Hyams (1983, 1986), Jaeggli and Hyams (1988), Lillo-Martin (1986), Guilfoyle (1984), Pierce (1992), Lebeaux (1987), and others have argued that the dropping of lexical subjects has its roots in the child's developing grammatical system. Specifically, it is argued that the child's grammar of English differs from the adult grammar in allowing null subjects. Hyams has proposed that this reflects an initial setting of the null subject parameter and is thus related to the typological variation found among adult languages. Others, notably Guilfoyle (1984), Borer and Wexler (1992), Radford (1990), and Guilfoyle and Noonan (1988), have related subject omission to other aspects of early grammar, for example, the absence of the Case Filter or functional categories, or (in the case of Borer and Wexler) the relaxation of an early requirement that each verbal element have a unique subject.

Although these analyses differ in theoretical and empirical detail, they have in common the attempt to achieve some explanatory power by predicting the simultaneous or

A much shorter version of this article was presented at the 1989 Boston University Conference on Language Development. We would like to thank the participants for the lively discussion that ensued. We would also like to thank Bob Berwick, Lois Bloom, Stephen Crain, Fernanda Ferreira, Lyn Frazier, Vicki Fromkin, Merrill Garrett, Jane Grimshaw, Yosi Grodzinsky, Maryellen MacDonald, and Virginia Valian, all of whom have contributed—many in obvious ways—to this article. We are also grateful to the two anonymous reviewers for their detailed, extensive comments and helpful suggestions; the article is much improved as a result of their efforts. All errors and misinterpretations remain our own. Our appreciation also to Sue Ingles of the Biomath Department at UCLA for the statistical analyses; to Sigga Sigurjónsdóttir and Tom Cornell for their help with the child language data; to Susan Hirsh and Jeannette Schaeffer for editorial assistance; and to Sandro Duranti for generously providing us with the Italian transcript data discussed in section 3.3.

Linguistic Inquiry, Volume 24, Number 3, Summer 1993
421–459

sequential development of particular linguistic structures. They also share the goal of providing a principled description of the intermediate stages of development by relating them to adult grammars, and theories of Universal Grammar more generally.

These grammatical accounts of subject omission stand in contrast to certain "performance" accounts. A number of researchers, for example, L. Bloom (1970), Pinker (1984), Valian (1991), and P. Bloom (1990), have argued that children acquiring non-null subject languages, such as English, drop subjects because of processing constraints. Although these authors do not provide a formal model, they propose that there is an upper bound on the length of utterance a young child can produce; hence, certain elements that may be grammatically represented fail to surface.

It has also been proposed that the dropping of grammatical subjects follows from a Principle of Informativeness (Greenfield and Smith 1976), according to which children may omit from their utterances that information which is most easily recoverable from context. On this account, children tend to utter those elements that carry the greatest informational content, independent of grammatical structure. Since subjects are often "old" or "given" information, they may be omitted.[1] Presumably, processing limitations are also what force children to limit their utterances to the most informative elements. The Informativeness account shares with the processing account the claim that a subjectless sentence is not a grammatical option for the child, and that the omission is due to some aspect of performance.

In this article we will evaluate these performance accounts of subject omission in child language, focusing primarily on the processing account. We will formalize these (rather sketchy) proposals and test the models against the statistical properties of child language data. As we will show, these models are incompatible with the basic statistical properties associated with subject omission in child language. We will also discuss a number of serious theoretical and conceptual problems with the particular processing account proposed by P. Bloom (1990). We will show that the grammatical approach to null subjects provides the tightest fit with the statistical data we examine, and the most explanatory account of the phenomenon.

In section 1 we present the theoretical and conceptual arguments in favor of the grammatical, principle-based account. In sections 2 and 3 we turn to the main point of the article, which is to assess the performance analyses in terms of the statistical properties of children's language.

[1] It should be noted that many of the rule-based analyses of subject omission cited in the text also assume that there is a pragmatic component to pro drop. For example, Hyams (1986) assumes that in the child's language, as in adult null subject languages, there are various pragmatic constraints on when subjects can be dropped, for example, when the subject refers to an established discourse topic. Thus, pragmatic and structural factors intersect to determine when a lexical subject will be used. Nonstructural pragmatic accounts, such as that proposed by Greenfield and Smith, differ in that they typically assume that the child's early language, or at least the phenomenon in question, is governed *uniquely* by pragmatic considerations. Crucially, on the pragmatic account subjectless sentences are *not* a grammatical option for the child. In section 2 we will consider the structural and nonstructural pragmatic accounts in more detail.

1 A Principle-/Rule-Based Approach to Child Language

Traditional studies of language acquisition have often found it difficult to establish a concept of explanation. The one clear fact was that the older children become, the more adult-like their language. Studies showed this generalization with respect to a wide variety of linguistic phenomena. To say the least, such a result proved to be not very satisfying from an explanatory viewpoint. One reason that it was difficult to go beyond such a result was that very little was assumed with respect to adult linguistic competence. In fact, it was often argued that child language should be studied without considering the adult system (Bowerman 1973). As a result, very little could be predicted about the actual course of development, and there was no principled relationship between one acquisition stage and the next, or between the intermediate stages and the target grammar.

In recent years, however, a number of scholars have attempted to study language acquisition within the principles-and-parameters framework of linguistic theory (Chomsky 1981). The idea is that a number of universal linguistic principles are innate in the child, and that the child must only learn the values of parameters and a lexicon for a particular language. Such a tightly constrained theory has had significant explanatory consequences for linguistics itself, and language acquisition theory aims to capture the same kind of explanatory power.

In this regard, one particular consequence of the framework has become quite important. Linguistic theory has attempted to show that a wide range of apparently unrelated phenomena in a particular language actually follow from the setting of a single parameter. This demonstration provides the theory with a certain amount of deductive strength. The same methodology has been extended to the study of actual development. In particular, language acquisition theory has attempted to show that a variety of apparently unrelated phenomena develop at the same time, perhaps as the consequence of the setting of a particular parameter. If it can be shown that just the right structures develop at the same time, or in a particular, relevant sequence, then language acquisition theory will also have achieved a certain deductive power.

One of the most extensively studied linguistic parameters is the null subject parameter, which captures the variation that languages exhibit with respect to whether the subject of a sentence must be phonologically realized or not. For example, in Italian (and Spanish) the sentence (2a) is fully grammatical, but in English the corresponding example (2b) is not.

(2) a. Lavorano molto in questa città.
    b. *(They) work a lot in this city.

Thus, in certain languages null subjects are a grammatical option, and there are a variety of discourse, pragmatic, and grammatical factors that determine when a subject will be realized as null. By contrast, in a language like English null subjects are not a grammatical option, and hence there is no situation in which a sentence like (2b) is acceptable.[2]

Given the principles-and-parameters framework, linguistic theory leads us to *expect* that children might avail themselves of a grammatical option presented by the null subject parameter even if it is not exhibited in the adult language of their community. Even more interesting is the fact that linguistic theory predicts that if children adopt the null subject option, they will also exhibit a number of other particular structures. As a case in point, it has been observed in a number of typologically diverse languages that there is a systematic association between the child's use of null subjects and various properties of the early inflectional system. Clahsen (1986), for example, has observed that for German-speaking children the end of the null subject stage coincides with the acquisition of an adult-like inflectional system and the systematic movement of the tensed verb to second position. This can be explained by a grammatical analysis of the sort proposed by Jaeggli and Hyams (1988), in which the movement of the verb to Comp (2nd position) blocks identification of the null subject, an analysis that is independently motivated by the facts of adult German (see Jaeggli and Safir 1989 for discussion).

Meisel (1987) reports a similar inverse relationship between null subjects and the development of inflectional morphology for German/French bilinguals. Correlations between the child's use of subjects and various properties of the inflectional system have also been noted in the acquisition of American Sign Language (Lillo-Martin 1986), Mauritian Creole (Adone and Verrips 1989), and Dutch (Weverink 1989).[3]

Pierce (1992) has proposed that the null subject phenomenon is related to yet other properties of early child language. Following Kitagawa (1986), Koopman and Sportiche (1991), and others, Pierce assumes that languages that take VP-internal subjects at S-Structure license null subjects. She argues on the basis of child language data in French, English, and Spanish that children initially have VP-internal subjects at S-Structure. Evidence for her hypothesis includes the fact that children use sentence-external negation, the fact that English-speaking children inflect auxiliaries earlier than main verbs, and the striking observation that French-speaking children systematically adopt post-

[2] For the present we abstract away from cases like *Wanna eat now?, Seems like it'll rain*—the restricted cases in which a subject may be omitted in English. Whether such cases are related to the null subject parameter is an interesting question, but one that goes beyond the scope of this article (but see Hyams 1992). More relevant to our present concerns is whether such sentences in the input could mislead the English-speaking child into believing that English is a null subject language, as Valian (1991) has suggested. The empirical fact that English-speaking children do not persist with a null subject grammar into adulthood suggests that the simple presence or absence of null subjects in the input is not what sets the null subject parameter, and that the trigger(s) is/are some other property of the language related to the null subject option. See Hyams 1986 for discussion of the triggering issue and null subjects, and see Gibson and Wexler 1992 for a discussion of the theory of triggers.

[3] Interestingly, Bloom, Miller, and Hood (1975) note that modals typically fail to occur in null subject sentences, once again suggesting a relationship between null subjects and Infl. This observation is distinct from that presented in Hyams 1986, where it was claimed that modals fail to occur altogether during the null subject stage. Rather, it seems that the effect is more specific; the child at this stage may have modals, but uses them only with overt subjects. The exclusion of modals from null subject sentences in child language has also been noted by Luigi Rizzi (personal communication).

verbal subjects. She also finds a correlation, discussed above, between the loss of null subjects and the development of inflection in French. Pierce's grammatical analysis, like many other studies noted above, relates the null subject phenomenon to a number of other properties of early child language and accounts for the range of phenomena in a principled fashion—that is, in a way linguistic theory would predict. The assumption is that the child is developing an adult grammatical system and that, depending on the precise structure of the system and the various interrelations that exist, each development may generate a wide range of effects.

Thus, there is a range of empirical data that any adequate account of subject omission in child language must explain. Central among these are (a) the relationship of null subjects to other aspects of grammar (e.g., the development of inflection, verb second (V/2) in Germanic, sentence-external negation, postverbal subjects in French) and (b) the various statistical properties associated with null subjects in child language (e.g., the significant difference that exists in the rate of subject vs. object omission). Any hypothesis concerning subject omission in child language must be evaluated on the basis of its success in handling both the statistical and the theoretical data, as well as its contribution to solving the explanatory problem of the learnability of language.

In this regard, it is important to emphasize that the processing and other performance accounts that we will consider do not address the important explanatory questions, questions with which the grammatical approach has had some success. As we will show, these analyses offer no independently established theoretical reason to expect that children will systematically omit subjects. Instead, reasons must be invented given the empirical observation. Moreover, these analyses say nothing about the correlations in development that are predicted by the grammatical accounts.

We have discussed a number of theoretical and conceptual advantages to the grammatical approach to subject omission in child language. The major point of this article, however, is to assess the performance accounts of the missing subject phenomenon in terms of the statistical properties of children's language. As we will show, the grammatical model fares better even on these limited empirical grounds. We will demonstrate that whatever performance effects may exist, they do not account for the range of phenomena associated with subject omission and hence do not vitiate the need for a principled structural analysis. The data we present are based on a detailed analysis of the transcripts of Adam and Eve (Brown 1973) from the CHILDES data base (MacWhinney and Snow 1985). Adam's speech was studied from eight 2-hour samples taken from the ages of 2;5 to 3;0 and Eve's speech from eight samples taken from the ages of 1;6 to 2;1.

2 Statistical Properties Associated with Missing Subjects

2.1 *Subject-Object Asymmetry*

One of the most striking facts about the dropping of arguments in early child language is that it is largely restricted to subjects. This asymmetry was discussed in Hyams 1983, 1986, but those works included no figures to support the claim, which was made on the basis of published data and descriptions. For the present study we calculated the rate of subject versus object drop in the CHILDES transcripts of Adam and Eve. We used the following criteria to determine missing subjects and objects. In calculating null subjects, we considered only those utterances that contained a verb, excluding the copula and auxiliary *be*. We excluded from the count imperative sentences and sentences that would be acceptable in colloquial English, for example, *Wanna go?, Don't know*. In calculating the rate of object drop, we considered all obligatorily transitive verbs. In addition, we counted any case where an optionally transitive verb was missing an object in a context where we deemed it necessary or appropriate.[4] For the purposes of this analysis, the data were divided into two periods, the early period being more squarely within the null subject stage. Each period represents four samples.

Table 1 shows the proportions of missing subjects and missing objects during the two periods.[5]

Table 1
Proportion of missing subjects and objects

	Adam	*95% CI*	*Eve*	*95% CI*	*Total*	*95% CI*
PERIOD 1						
Subjects	55%	(52–59)	39%	(35–42)	48%	(45–50)
Objects	7%	(5–9)	13%	(10–17)	9%	(7–11)
PERIOD 2						
Subjects	29%	(26–32)	15%	(12–17)	22%	(20–24)
Objects	11%	(9–13)	4%	(2–5)	8%	(6–9)

The results show a very large difference in the rate of missing subjects and objects for both children.[6] During period 1 Adam is dropping subjects at a rate of 55%, whereas objects fail to appear only 7% of the time in obligatory contexts. During period 2 the

[4] Notice that this procedure probably overestimates the amount of object drop since the child may sometimes intend the intransitive usage of the optionally transitive verb.

[5] P. Bloom (1990) also calculates the proportion of missing subjects and objects. His results for Adam are roughly equal to our Adam period 1, 57% subject omission versus 8% object omission. His figures for Eve, however, are very different from our own. His results show an even stronger subject-object asymmetry than we obtained, 61% subject omission versus 7% object omission. The large asymmetry in the proportion of missing subjects and objects is also amply documented by Bloom, Miller, and Hood (1985) for the four children they studied, Eric, Gia, Kathryn, and Peter.

[6] We have provided the confidence intervals (CI) for each proportion in table 1. If two 95% confidence intervals do not overlap, then the difference in proportions is significant at the .05 level. For example, for Adam period 1, the CI for subjects is 52–59, whereas the CI for objects is 5–9. Each of the differences in table 1 is significant at the .05 level.

rate of subject drop declines to 29%, whereas object drop increases slightly to 11%. Eve drops subjects at a lower rate than Adam. Nevertheless, there is still significantly more omission of subjects than objects, 39% versus 13% at period 1 and 15% versus 4% at period 2.

If we take 90% correct as the criterion for acquisition, which is standard (Brown 1973), these results show that children respect subcategorization requirements of transitive verbs as early as period 1, though they often appear to violate the requirement of an obligatory subject (about 50% of the time). We take this subject-object asymmetry to be a central empirical fact to be accounted for in addressing the question of why children omit subjects. Most grammatical analyses predict this asymmetry. Children omit lexical subjects, but rarely objects, because null subjects are a grammatical option for the child.

Let us now consider how this striking asymmetry is handled by the various performance analyses. In the following section we discuss the Informativeness approach and in section 3 the processing approach.

2.2 An Informativeness Account of the Subject-Object Asymmetry

As noted earlier, Greenfield and Smith (1976) propose that in describing an event, children tend to omit from a sentence that aspect of the situation which is "taken for granted," "presupposed," "less informative," and so on. According to them, "this explains the frequent omission of the subject in speech around the 3-word level" (p. 223). As discussed in Hyams 1986, this explanation is not particularly satisfying. Children at this stage talk very much in the here and now. Thus, there is a sense in which most of what they say is specific, that is, assumed vis-à-vis discourse.[7] When Kathryn (Bloom 1970) says *Man making muffins* while looking at a picture of a man making muffins, each element in the sentence is given by context. Thus, by the Principle of Informativeness any one of these elements is expendable.

Greenfield and Smith get around this kind of objection by claiming that the child's perception of what is contextually salient, and hence likely to be omitted, may differ from the adult's. But since we have no insight into the child's perception of a situation aside from what can be inferred from her language, the Principle of Informativeness has no predictive or explanatory value. The circularity implicit in accounts of this sort is also noted by Bloom, Miller, and Hood (1975).

We may be able to derive an empirical prediction from an Informativeness model, however. Notice that if children do not have a null subject grammar, then the pragmatic function of a missing argument is essentially equivalent to the pragmatic function of a

[7] What Greenfield and Smith have in mind is that the information is assumed in discourse. The term for such information in the pragmatics literature is *specific* information, and we will adopt this term. Notice that "specific" information and "presupposed" information in a technical sense can be quite different.

pronoun. Both the missing argument and the pronoun serve to encode specific information.[8]

We will refer to both pronouns and null arguments as *specific arguments;* so pronominal and null subjects are *specific subjects* and pronominal and missing objects are *specific objects*. Assuming that children use pronouns during the period in which they drop arguments, we make the prediction in (3).

(3) All else being equal, the ratio of *missing subjects* to *specific subjects* is equal to the ratio of *missing objects* to *specific objects*.

Why should this be so? According to the Informativeness hypothesis, children have two mechanisms for encoding information that is "presupposed" or "taken for granted": they can pronominalize an NP or they can drop it. According to this hypothesis, then, an argument is just as likely to be dropped as pronominalized. However, let us suppose that there is another factor, perhaps having to do with processing capacity, that explains why in certain cases an argument is dropped rather than pronominalized. We can assume for the sake of discussion that missing arguments impose less of a processing load than overt ones, as P. Bloom (1990) and others have suggested, though we return to this point in section 3 and the Appendix. So there is a certain drop rate attributable to processing constraints. The Informativeness hypothesis, however, predicts that the drop rate should be constant across grammatical functions; that is, there should be no difference in the rate of omission for subjects versus objects. To the extent that there is a difference, it must be due to reasons other than those associated with Informativeness. This is the logic behind the prediction in (3).

In table 2 we list the proportions of missing, pronominal, and specific arguments used by Adam and Eve for the eight samples. Recall that *specific arguments* represent the sum of null and pronoun arguments. In table 3 we report the ratios of missing arguments to specific arguments in subject and object position.

On average, these children drop specific subjects about twice as often as they drop specific objects. Or, in other words, information that is "specific" ("taken for granted," "given," "redundant," etc.) is most often pronominalized when it occurs in object position, but dropped or pronominalized in about equal proportions when it occurs in subject position. This difference is not predicted by the Informativeness model.[9] Notice, however, that under a grammatical model such a result is expected since the option to drop a specific argument is available only for subjects. In summary, to the extent that the Informativeness model can be tested, it fails the empirical tests. The grammatical model, on the other hand, is consistent with the data.

[8] In a pro-drop language, such as Italian, the pragmatic function of null arguments differs from that of pronouns. Null subjects are used to refer to an established discourse topic or given information, whereas pronouns are typically used for contrast, emphasis, and so on, and to disambiguate antecedence.

[9] Table 3 shows the confidence intervals for all proportions. (See note 6 for explanation of confidence intervals.) The difference between the ratios of missing arguments to redundant arguments in subject versus object position for both Adam and Eve is significant at the .05 level.

Table 2
Proportions of missing, pronominal, and specific arguments

	TYPE OF ARGUMENT		
	Missing	*Pronominal*	*Specific*
ADAM			
Subjects	.41	.38	.79
Objects	.089	.33	.41
EVE			
Subjects	.26	.46	.72
Objects	.07	.2	.27

Table 3
Ratios of missing arguments to specific arguments

	Adam	*95% CI*	*Eve*	*95% CI*
Subjects	.52	(.495,.547)	.36	(.330,.338)
Objects	.21	(.178,.246)	.27	(.216,.327)

2.3 A Topic-Drop Analysis

In the previous section we argued against a particular nonstructural pragmatic hypothesis concerning the basis for child null subjects. However, there is another sense of "pragmatics," one used in linguistic theory rather than in psychological theory. Here pragmatics is viewed as being based on mental principles, principles that might have their basis in UG and might interact with syntactic and semantic principles. In fact, there is evidence that pragmatic principles of this sort are part of the licensing conditions for null subjects in some languages. For example, Dutch is a V/2 language. Any constituent can appear in first position, which is generally the topic position (SpecCP). This topic can be dropped. For example, (4b) is a reply to (4a) (both from Schaeffer 1990). In (4b) the topic (the object, perhaps the Dutch equivalent of *that* or *it,* that is, *Rainman*) is dropped.

(4) a. Ga je mee naar *Rainman* vanavond?
go you to *Rainman* tonight

b. Heb ik al gezien.
have I already seen

That the dropped topic came from first position can be seen from the grammatical sentence in (5).

(5) *Rainman* heb ik al gezien.

The subject *ik* 'I' is not dropped in (5). It can only be dropped when it is in first—that is, topic—position. Thus, (6) is ungrammatical.

(6) **Rainman* heb al gezien.
Rainman have already seen

De Haan and Tuijnman (1988) propose that Dutch-speaking children have a process of topic drop. They found that children uttered sentences like (4) and (5), but not (6). That is, children had topics in first position and they often dropped the constituent in first position whether it was a subject or an object. But they almost never dropped a subject that was not in first position.

Similarly, Wang et al. (1992) observe that Chinese-speaking children also drop (subject and object) topics, as is permissible in that language (see Huang 1984). Given this, it is entirely possible that missing subjects in early English are also the result of topic drop (as suggested in Jaeggli and Hyams 1988) and not pro drop per se.[10] We would then have to explain why the dropping is restricted to subjects in English, but affects both subjects and objects in Dutch and Chinese.

A possible pragmatic account of the cross-linguistic differences in the acquisition data goes as follows. In order to be dropped, a constituent must be outside the "nuclear scope" in the sense of Diesing (1988) and Kratzer (1989).[11] We will not attempt to develop this account in great detail, but essentially it means that to be dropped, a constituent must be scoped outside the VP. In Dutch any constituent may wind up in SpecCP and thus may be omitted under topic drop. There is good evidence that young children (in the null subject stage) learning V/2 languages like Dutch and German topicalize constituents appropriately, moving them to SpecCP.[12] Similarly, in Chinese constituents topicalize readily, and we assume that young children know this process of topicalization. English, however, has a much less robust pattern of topicalization. Suppose that English-speaking children at this age have no topicalization process.[13] The subject is already

[10] The topic-drop account of early English proposed by Jaeggli and Hyams (1988) differs in technical details from the one about to be proposed. See also Hyams 1992.

[11] We thank an anonymous reviewer for insightfully suggesting an account along the following lines, though our proposal differs from it in many details.

[12] For discussion, see Hyams 1992, Poeppel and Wexler 1991, Weissenborn 1990, and Wexler 1991.

[13] In unpublished work Chien and Wexler (personal communication) present evidence that even older English-speaking children don't have topicalization. It seems reasonable to assume that early topicalization in Dutch is related to that language's V/2 nature (which children in the pro-drop stage know; see De Haan and Tuijnman 1988, Poeppel and Wexler 1991, 1993). English-speaking children develop the limited English V/2 (e.g., inversion in interrogatives) later than children learning full V/2 languages, and their later use of topicalization may be related to this fact.

outside the nuclear scope, so it can be dropped in any event. The object, however, must be scoped out by topicalization; only in this way can it be dropped. Thus, on the assumption that English-speaking children do not have topicalization, objects will not drop. It would follow on this account that only subjects may drop in English, whereas both subjects and objects would be affected in Dutch and Chinese.

The topic-drop idea may also account for certain statistical differences between English- and Italian-speaking children. Valian (1991) notes that young Italian-speaking children drop subjects at a rate of about 70%, whereas the English-speaking children in her study at the same developmental level dropped subjects at a rate of only 30%.[14] Valian argues on this basis that the Italian-speaking children are dropping subjects as a result of a grammatical process of pro drop, whereas English-speaking children are dropping them as a result of performance constraints. However, there is no reason why performance constraints should yield a drop rate of 30% and grammatical factors a higher one; that is, there is no performance model from which these frequency effects would follow.

An alternative to the performance account of the English/Italian frequency differences is one that follows from the difference in grammar types discussed above. By our hypothesis, Italian-speaking children have a pro-drop grammar, in which null subjects are identified by "rich" Agr(eement) as in the adult language, whereas English-speaking children have a topic-drop grammar. Since every finite sentence contains Agr, null subjects will always be a grammatical option for the Italian-speaking child.[15] However, not every sentence has an appropriate (subject) topic, and the English-speaking child will therefore have fewer opportunities to drop subjects.

This account predicts that Chinese-speaking children should drop subjects at a lower rate than Italian-speaking children, all else being equal (and at a rate similar to that of English-speaking children). According to Wang et al. (1992), the Chinese-speaking children in their study dropped subjects at a rate of about 56%, hence less often than Valian's Italian-speaking children. Moreover, Wang et al. report that the difference in the rate of subject omission between their Chinese- and English-speaking subjects at the earliest developmental stage was not statistically significant. Thus, the English- and Chinese-speaking children pattern alike with respect to missing subjects (though they differ with respect to missing objects for the reasons outlined above).

Thus, our topic-drop account fits well with a wide range of empirical data, though

[14] Our own data, as well as P. Bloom's (1990), show a higher rate of subject omission for English-speaking children than Valian obtained. Our figures, given in table 1, show an average drop rate of 47% for Adam and Eve. Bloom's (1990) figures diverge even more sharply from Valian's. According to his calculations, Adam dropped subjects at a rate of 57% and Eve at a rate of 61%.

[15] For evidence that Italian-speaking children do have Agr(eement), see Hyams 1983, 1991, Schaeffer 1990, and Guasti 1992, and for evidence that agreement is acquired early in Romance and Germanic languages generally, see Wexler 1991. One empirical difference between Italian and English child language is the use of agreement morphology; Italian children acquire the inflectional affixes (at least for 1st, 2nd, 3rd person singular) by age 2. English-speaking children acquire the 3rd person *-s* ending much later. Thus, null subjects in early Italian would be identified the same way they are in the adult language.

space limitations prevent us from providing a more detailed analysis.[16] Topic drop is clearly a pragmatically governed principle. However, it is one that interacts directly with the syntax and semantics. Moreover, the account outlined here is constrained by principles of UG. In these respects, it differs markedly from the Informativeness account discussed earlier.

3 A Processing Account of Subject Omission

In this section we consider the idea that children omit subjects because of a processing overload of some sort. In sections 3.1 and 3.2 we first discuss apparent evidence in favor of this processing account and then outline general problems with it as it is presented in the literature. We discuss, in particular, the "VP length effect" obtained by Bloom (1990). In section 3.3 we show that this effect cannot be due to processing limitations, but is in fact further evidence for a grammatical null subject analysis. The various processing proposals we consider (e.g., Bloom 1990, Valian 1991) are rather inexplicit and as a result they are difficult to test. Thus, beginning in section 3.4 we present an explicit processing-overload model of subject omission, which we call the "Output Omission Model," and we test this model against the statistical data.[17] In section 3.4 we present the model intuitively; in section 3.5 we provide a formalization; and in section 3.6 we consider further consequences of the model.

3.1 Evidence Favoring the Processing Approach

L. Bloom (1970) was the first to propose that children drop major constituents of a sentence, such as subject, as a function of the underlying grammatical complexity of the sentence. She found, for example, that children's negative sentences lack subjects more frequently than their nonnegative sentences. On the assumption that negation adds complexity to the sentence, children, who operate under more severe processing constraints than adults, would be likely to drop some other element, according to Bloom. Feldman, Goldin-Meadow, and Gleitman (1978) report a similar kind of result in their study of the spontaneous sign language of isolated deaf children. They found that these children were more likely to omit subjects with transitive verbs than with intransitive verbs. These results suggest that the subject is more likely to be omitted when the VP is longer or when the argument structure of the verb is more complex.

However, other studies have had inconsistent results. Bloom, Miller, and Hood

[16] We do not consider how the topic-drop phenomenon relates to the adult language or how the child changes his analysis. In particular, the question arises how the English-speaking child's grammar relates to the more limited kind of topic-drop grammar in the adult. Why do English-speaking adults not drop as wide a range of subjects as children do? Also, when topicalization develops, why are objects not dropped in English and other languages? For an extension of our proposal that attempts to answer these questions, see Hyams 1992.

[17] In the Appendix we consider the OOM in conjunction with another processing model cited by Bloom (1990), the model proposed by Yngve (1960).

(1975) carried out an extensive study to determine the effect of grammatical complexity, as well as various lexical and discourse factors, on utterance length in child language. They considered a number of grammatical elements that might induce complexity, including negation, nominal and verbal inflection, prepositions, adverbs, possessives, verbal particles, definite articles, and demonstratives. Their results showed that, overall, grammatical complexity had a significant effect on utterance length for only two of four children, and for these two children the effect was significant in less than half the trials. Bloom, Miller, and Hood then performed a more fine-grained analysis, in which they looked at the effect of some of the individual elements on utterance length. This analysis showed that negation, verbal particles, and adjectives did correlate with utterance length. Thus, children were more likely to use negation with utterances containing two major constituents (e.g., *No eat it*) than with utterances containing three constituents (e.g., *I no eat it*). However, prepositions, determiners, double object constructions, and various forms of inflection did not covary with utterance length and in fact, even clausal subordination and coordination showed no significant effect on utterance length. Thus, according to the results of this study, a child is just as likely to say, for example, *Mommy want eat it* as *Want eat it,* or *Mommy give them milk* as *Give them milk.*

P. Bloom (1990) has also attempted to show processing effects on utterance length. Bloom analyzed the CHILDES transcripts of Adam, Eve, and Sarah, comparing VP length (in terms of number of words) in sentences with and without subjects, as well as in sentences with pronominal subjects. His results show that VP length increases as a function of subject type: null, pronominal, or lexical (i.e., full NP). The VP is shortest when the subject is a lexical NP, longer when the subject is pronominal, and longest when the subject is null. Bloom's results are shown in figure 1.

Crucially, Bloom assumes that because lexical subjects are on average phonetically longer than pronouns, they impose a heavier processing load and, similarly, that pronouns impose a heavier load than null subjects. The heavier the subject, the fewer processing resources available for the rest of the sentence and hence the shorter the VP.[18] According to Bloom, the VP length results support the processing model of missing subjects over the grammatical model.[19]

[18] Alternatively, the VP is planned first and the subject is either dropped or not depending on the resources taken up by the VP. Since Bloom does not present a specific processing model, it is difficult to know exactly how processing "overload" results in subject omission on his account. In sections 3.4 and 3.5 we will provide a formal processing model intended to capture these ideas.

[19] Bloom presents two other findings: first, that the proportion of pronouns to full NPs is greater in subject position than in object position, and second, that nonpronominal overt subjects are shorter on average than nonpronominal overt objects. Bloom argues that these two results further support a processing account.

With respect to the first point, it is likely that the same result would be found with English-speaking adults. Subjects tend to be topics, which are therefore more likely to be pronominalized. Thus, the effect is probably a pragmatic one.

The discussion around the second finding is somewhat misleading since the length measure here is number of words (which correlates with structural complexity) and not phonetic length, as in the analysis of VP length discussed in the text. In this case, as well, Bloom does not perform a similar analysis on adult controls and

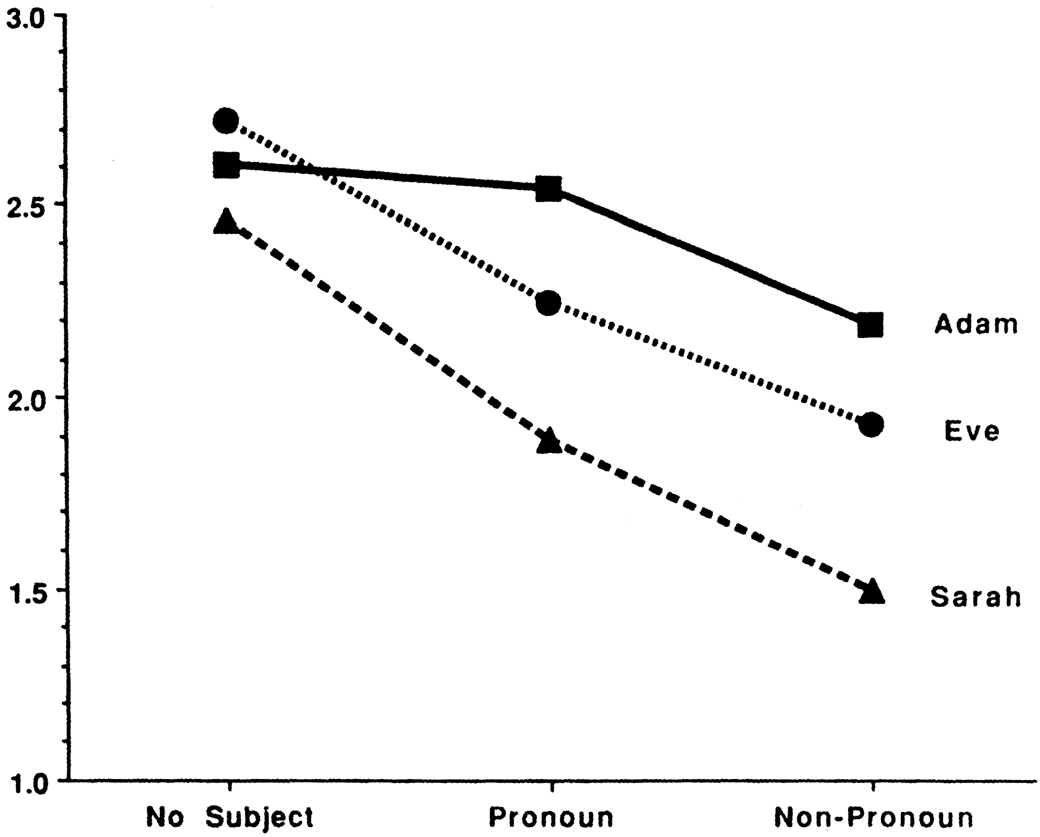


Figure 1
VP length as a function of subject size for three English-speaking children. (From P. Bloom 1990:fig. 1.)

3.2 Some Problems with the Processing Approach

In this section we outline some problems with the processing account in general, and with Bloom's analysis in particular.

First, as discussed in section 2, the major *statistical* fact about null subjects is that they far outnumber null objects and this central phenomenon is not accounted for or

---

hence it is impossible to conclude anything about the children's behavior. If we found, for example, that the same result holds for adult English speakers (which again seems likely), then we would need to ask why adult speakers do not drop subjects as well, assuming, as Bloom does, that missing subjects and "short" subjects are due to the same processing constraints.

The importance of using adult controls in these analyses will become apparent in section 3.3, when we compare Bloom's statistical data to data from Italian-speaking adults.

predicted in any way by the various processing accounts (Bloom 1990, Valian 1991, among others). The processing account allows only that some element(s) will be omitted as a function of increased complexity. Bloom (1990) does attempt an explanation for the subject-object asymmetry, but, as we will discuss shortly, his proposals receive little empirical or theoretical support.

The second problem relates to the central claim of the processing hypothesis, namely, that there is an upper bound on the length of utterances a child can produce. Under such an account, it would be odd if children who are dropping subjects were also systematically producing longer utterances. Hyams (1983, 1986) notes that children who produce subjectless sentences are clearly capable of producing longer utterances. Bloom (1990) disputes this claim. He states that during the missing subject stage, children produce longer sentences only "occasionally" and that they have "a tendency to produce short utterances" (p. 493). Although it is difficult to know exactly what is meant by "occasional" and "tendency," the statistical data reported by Bloom, Miller, and Hood (1975) do not support these claims under any reasonable construal and present quite a different picture of the facts.

If we look at the last data point for each of the four children examined by Bloom, Miller, and Hood (1975), we see that two of them, Eric and Kathryn, are using two- and three-term utterances in roughly equal proportions. Thus, at time (stage) VI, 46% of Eric's utterances containing action verbs are verb-object utterances, whereas 41% are subject-verb-object; Kathryn at time III produces 46% verb-object utterances as compared to 48% subject-verb-object. Peter at time IX produces 54% verb-object utterances and 34% subject-verb-object. Here the difference is larger than for Eric and Kathryn; however, 34% hardly seems to qualify as "occasional." Gia at time V produces 28% verb-object utterances and 61% subject-verb-object. This stage of Gia's development is difficult to evaluate since she may be on her way out of the null subject stage. However, if we look at the previous stage, Gia IV, we find the same pattern that the other children produce: verb-object and subject-verb-object in roughly equal proportions, 37% and 30%, respectively.

These data fully support Hyams's original claim that children who drop subjects are able to produce longer utterances and do so regularly.[20] Thus, if children are dropping subjects because of a processing bottleneck, why do the computational difficulties arise in certain instances, but not in others? A processing analysis of the phenomenon should tell us what the variables are that affect the children's output.

Let us now consider how the subject-object asymmetry is handled on the processing account. In order to account for the significant difference in the rate of subject omission over object omission, Bloom (1990) proposes that the beginning of the sentence imposes

[20] In fact, Bloom, Miller, and Hood (1975) make the point directly when they say that over the period they studied "three- and four-constituent utterances increased, but the two-constituent utterances [i.e., subjectless sentences/NH & KW] continued to occur" (p. 10).

a heavier processing load than the end of the sentence.[21] In current processing models, however, processing complexity is typically related to structural properties of sentences and not serial position.[22] (For a general overview, see Frazier 1985.) Moreover, there is little in the way of empirical evidence that the beginning of the sentence is harder to process. In fact, as far as children are concerned, empirical data from comprehension and production suggest just the opposite.

With respect to comprehension, for example, Newport, Gleitman, and Gleitman (1977), in their important "Motherese" study, show that the rate of acquisition of auxiliaries correlates with the frequency of yes-no questions in the input, where the auxiliary is in initial position. They argue on the basis of this result that the beginning of the

[21] Bloom cites Yngve (1960) in connection with his claim that subjects require more processing load than objects, or alternatively, that the beginning of the sentence is more difficult to process than the end. We discuss Yngve's model in the Appendix. Bloom also cites Bever (1970) as providing evidence that the beginning of the sentence is harder to process than the end. In addition to the fact that Bever is concerned with comprehension, not production, we have found nothing in his paper that suggests or supports this claim. Bever's paper deals with the possible perceptual underpinnings for various grammatical structures. What Bever does point out, which may be the source of Bloom's claim, is that the interruption of major constituents (e.g., subject NPs or VPs) by relative clauses, heavy adverbial, adjectival, or prepositional phrases, and so on, induces perceptual complexity and that *in such instances* the complexity is reduced when the intervening material is moved to the right, essentially to the right periphery of the major constituent (Bever's Principle G; pp. 330ff.). Some of Bever's examples are given in (i)–(v), where the second member of each pair is judged less acceptable than the first.

(i) a. John called up the girl in the white dress.
b. ?John called the girl in the white dress up.
(ii) a. John showed the girl the book that I liked a lot.
b. ?John showed the book that I liked a lot to the girl.
(iii) a. John walked briskly in a slightly northerly direction.
b. ?John walked in a slightly northerly direction briskly.
(iv) a. The steel and artificially strengthened fibre plastic tube broke.
b. ?The artificially strengthened fibre plastic and steel tube broke.
(v) a. The machine is bulky and incredibly hard to operate without the appearance of at least one malfunction.
b. ?The machine is incredibly hard to operate without the appearance of at least one malfunction and bulky.

At first glance some of these examples (i.e., the contrasts in (i)–(ii)) might be construed as support for the notion that processing capacity increases as the sentence moves from left to right, and hence that listeners are better able to handle complexity later on in the sentence. However, the sentences in (iii)–(v) show that it is not simply the serial position of a phrase in a sentence that accounts for perceptual difficulties, but rather its relationship to the other elements within the same major constituent. For example, the problem with (ivb) is not that there is a heavy AP at the beginning of the sentence, as the perfectly acceptable (vi) shows.

(vi) The artificially strengthened fibre plastic tube broke.

Rather, in (vb) it is the relative weight of the first AP and lightness of the second AP *(bulky)* that creates a degree of unacceptability. When the two APs are balanced, even though they may both be heavy, the perceptual problem disappears, as in (vii).

(vii) The artificially strengthened fibre plastic and double-walled reinforced steel tube broke.

Moreover, the effect holds within both subjects and predicates (see (iv)–(v)). Thus, whatever accounts for the perceptual effects discussed by Bever, it is not simply the position of the phrase within a sentence, contrary to what Bloom implies.

[22] More generally, the classical result in studies of short-term memory is that the beginning and end of an unstructured list of items are both easier to recall than the middle. This is the so-called bowed serial position curve. Of course, there is no reason that this process should have anything to do with sentence production.

sentence is perceptually salient for children. Similarly, Curtiss and Tallal (1991), using an error analysis of children's performance on picture identification tasks, show that children attend more reliably to material in the first half of the sentence.

Although there are even fewer experimental studies that focus on production abilities in young children, the empirical data that do exist also fail to support the hypothesis that the beginning of the sentence exerts a more severe processing load. Ferreira and Morrison (1990) note that in an elicited production study, the most common error children make with long subjects such as *the tall man* is to omit the head Noun, leaving behind *the tall*. Thus, if children divide the sentence into two production units, one containing the subject and the other containing the VP, as Ferreira (1991) argues is the case for adults, then in fact these children are showing a tendency to drop material toward the end of the production unit, directly counter to Bloom's proposal.

With respect to adult production, there are some effects that suggest an increased processing load clause-initially. For example, speakers tend to exhibit longer pauses/hesitations clause-initially. However, as Merrill Garrett (personal communication) points out, that could be due to "consolidation" effects at the end of the preceding sentence as much as from "planning" for the upcoming sentence. Similarly, Fernanda Ferreira (personal communication) notes that the pause duration would indicate that the load is high *before* articulation begins, not that the load is high at the beginning of the sentence.[23]

But even if the pause/hesitation effects were due to an increased processing load, we would still need to ask why an increased processing load manifests itself as *hesitation* in adults but as *subject deletion* in children. It is not at all obvious why this should be so. Why don't children show the same kind of hesitation effect as adults—that is, why don't they include the subject, but hesitate before or while saying it?[24]

The empirical evidence for an increased processing load sentence-initially in adults is, in the best case, equivocal. For children, the only available evidence in fact contradicts this assumption. Thus, the processing hypothesis still leaves unexplained the most salient fact about subject omission in child language, namely, that it is largely restricted to subjects.[25]

[23] Merrill Garrett and Victoria Fromkin (personal communication) both note that there are no serial position effects in their speech error data; that is, speakers do not tend to make more speech errors at the beginning of a sentence, which would be predicted if there were indeed a greater processing load clause-initially.

[24] We are grateful to Merrill Garrett for posing this question to us.

[25] More recently, Gerken (1991) proposes a performance constraint defined on the metrical structure of children's utterances. Her hypothesis is that children omit from their productions the weak syllable in an iambic (i.e., weak-strong) foot, but not in a trochaic (strong-weak) foot. This proposal captures the subject-object asymmetry insofar as subject pronouns that are sentence-initial form an iambic foot with a strongly stressed verb, as in (i), whereas an object pronoun can be the weak syllable of a trochaic foot, as in (ii), and hence would not delete.

(i) she KISSED + the DOG
(ii) the DOG + KISSED her

Gerken's hypothesis also predicts that children will omit articles from iambic feet (e.g., the object article in (i) and the subject article in (ii)), more frequently than from trochaic feet (e.g., the object article in (iii)).

We turn finally to the result, obtained by Bloom (1990), that VP length varies as a function of subject type. The hypothesis that the relationship between subject type (NP, pronoun, or null) and the length of the VP is based on processing factors, as Bloom maintains, is problematic in several respects. First, there is no evidence from any other aspect of child language that phonetic length determines production or complexity. In fact, in studies of language acquisition a rough measure of complexity, mean length of utterance (MLU), is usually considered to be number of morphemes (Brown 1973) and not phonetic length.[26]

Moreover, it seems unlikely that a full lexical subject, such as *Eve,* would be more difficult to process/produce than a pronoun, given that a pronoun involves at least one additional processing operation, that of accessing its antecedent. In fact, Ferreira and Morrison (1990) report experimental evidence showing that children do indeed have greater difficulty with pronominal subjects than with lexical subjects. The children in this study were asked to repeat some part of a sentence, for example, the subject. Four different subject types were presented—pronoun, single Noun (name), Determiner + Noun, and Determiner + Adjective + Noun—in sentences like *He/Dick/The farmer/ The silly farmer ate supper with the pigs.* Ferreira and Morrison's results show that children are much less successful with pronominal subjects (56% correct) than with names (79% correct). The latter percentage refers to one-syllable names, such as *John.* With two-syllable names the children perform even better (86% correct), and in fact, they do best with full Det + N subjects, such as *the farmer* (88% correct). They perform least well on NPs containing adjectives, such as *the silly farmer* (41% correct), but as noted earlier, the only omission error they make with this subject type is to drop the

---

(iii) PETE + KISSED the + DOG

The results of Gerken's elicited imitation task bear out these predictions.

Gerken's results are certainly suggestive, and represent a more principled approach to subject omission than the other performance-based accounts discussed in the text. However, they raise questions of both an empirical and a theoretical nature. First, English-speaking children also omit subjects that are not in initial position and that do not necessarily form an iambic foot with the verb, for example in *wh*-questions (though subject omission in *wh*-questions is rarer than in declaratives; Rizzi 1992) and in negatives, as in (iv) and (v).

(iv) What doing? [WHAT he + DOing]

(v) No want stand head. [NO he + WANT . . .]

Negatives such as (v) occur during the period of "sentence-external negation" (Bellugi 1967, Pierce 1992); hence, the omitted pronoun would follow the negation and form a trochaic foot with it, as would be the case for *wh*-questions. Also, Gerken's analysis is at odds with Ferreira and Morrison's (1990) result discussed in section 3.2, in which the most common error children made in an elicited imitation task was to omit the head Noun in subjects such as *the tall man.* Gerken predicts that the determiner should be dropped in these cases and not the head Noun.

Theoretically, Gerken's analysis comes up against the same question as the other performance-based accounts, namely, Why does the child deal with complexity by omitting constituents, or on this analysis, unstressed syllables? Why don't we simply find hesitation effects, as we do with adults? (See section 3.2.) Moreover, why does this happen in iambic feet, and not trochaic feet? Even more significantly, null subjects seem to be correlated with a number of other grammatical properties, for example, "optional infinitives" (Wexler 1991, 1993). Yet empirical results in a number of languages show that these properties (e.g., optional infinitives) do not reflect any kind of "dropped" constituent.

[26] We are grateful to Yosef Grodzinsky for reminding us of this fact.

head Noun. Ferreira and Morrison's results run directly counter to Bloom's claim that longer subjects are more difficult to produce. It seems unlikely, then, that the VP length effects obtained by Bloom are related to processing difficulties associated with phonetically longer subjects.

3.3 VP Length Effects in Null Subject Languages

Suppose, on the other hand, that the VP length effect is a property related to the use of null subject languages. If a child is speaking a null subject language, as the grammatical account assumes, we would expect the child's utterances to show the VP length effect. The crucial test of this hypothesis would be to determine whether the VP length effect holds in a known null subject language. That is, in a known null subject language, does VP length vary depending on whether the subject is null, pronominal, or lexical, as it does for the English-speaking child?

We have analyzed the spontaneous discourse of several Italian-speaking adults, in which we compared type of subject and length of VP, as Bloom did with English-speaking children. We examined five transcripts. The taping sessions ranged in length from 7 to 30 minutes. The subject of each sentence was coded as null, pronominal, or lexical. Sentences with postverbal subjects were excluded, as were sentences containing relative clauses, clefts, existential *ci* 'there', *wh*-phrases, and imperatives. In calculating VP length, we included the verb and its arguments (optional and obligatory) and inflectional elements such as negation and auxiliaries. We excluded right-dislocated and left-dislocated constituents, as well as interjections. Repetitions were not counted. The results of our analysis are given in figure 2.

As figure 2 shows, exactly the same trend was found for Italian-speaking adults that was found for English-speaking children (cf. figure 1). A 2-way analysis of variance (ANOVA) found that the mean VP lengths for the three subject types are significantly different ($p = .0001$). A Tukey multiple comparisons test showed that the means for null versus lexical subjects are significantly different at the .05 level. The other pairs of means (i.e., null vs. pronoun, and pronoun vs. lexical NP) are not significantly different.[27] Finally, a test for linear trend also gave a significant result ($p = .0001$). Thus, the results from Italian-speaking adults are directly comparable to those from English-speaking children.[28]

If we assume that the relationship between VP length and subject type in child language is a processing effect, what do we say about the Italian-speaking adults? If we were to argue that adults drop subjects because of processing reasons and not because this represents a grammatical option, then a vast array of linguistic results could not be accounted for. The VP length effect is true for Italian-speaking adults, and thus possibly

[27] Bloom (1990) did not perform a significance test on these pairs of means.

[28] Bloom did separate analyses on the three children. This was not possible for the Italian-speaking adults since they varied from transcript to transcript.

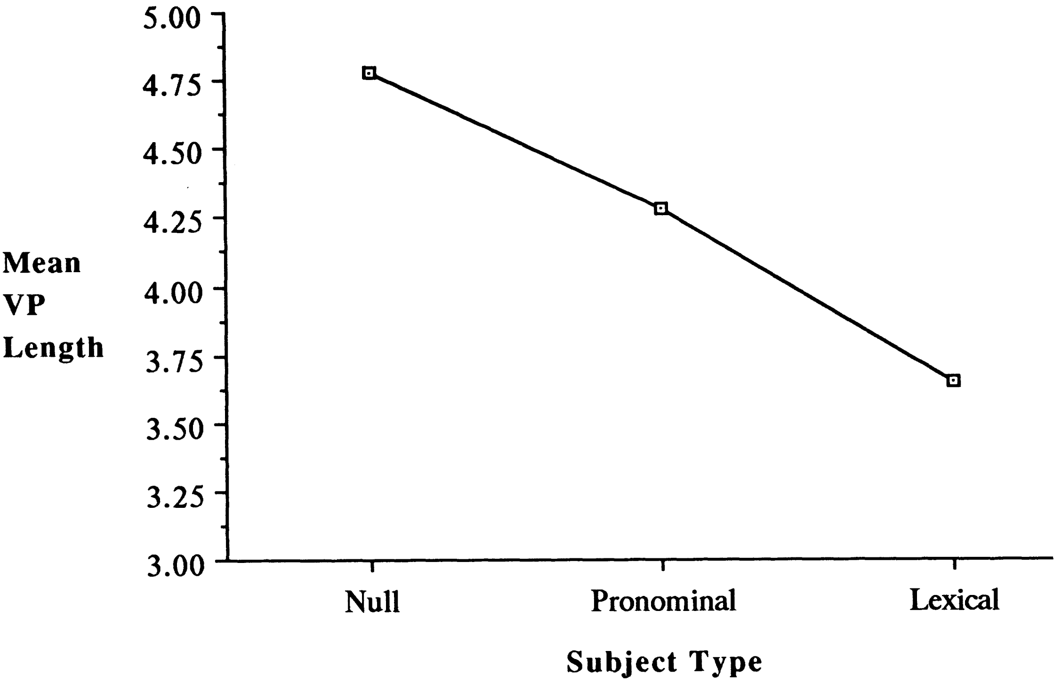


Figure 2
VP length as a function of subject type for speakers of Italian

true for null subject languages in general. Bloom's results show that English-speaking children in the subject omission stage also exhibit the VP length effect. Thus, a property of null subject languages turns out to be a property of child language in the subject drop stage. We would argue that, far from providing evidence of a processing limitation in children, the VP length effect is yet another reason to believe that English-speaking children are speaking a null subject language.[29]

[29] Perhaps the VP length effect results from an interaction of pragmatic and grammatical factors. As is well known, nonobligatory constituents (e.g., adjuncts, and certain objects) can be dropped from the sentence. This is more likely to happen if the constituent contains presupposed or old information. Consider the extreme case of null subjects versus names in a null subject language. (Similar effects might hold for indefinite subjects.) A null subject is typically presupposed; a name more often constitutes new information. Therefore, the predicate in a null subject sentence more often contains new information than the predicate in a sentence whose subject is a name. Given that old information can be elided from the predicate if it is not grammatically obligatory, these processes will result in longer VPs in the null subject sentences. Of course, this effect is only statistical. There is a second grammatical process that might contribute to the VP length effect. Even if old information in the VP is not elided, it is likely to be pronominalized, and this can happen even with obligatory constituents. Pronominalization would also shorten the VP by eliminating determiners and modifiers.

We believe that a similar kind of explanation will account for the intermediate status of VP length in sentences with pronominal subjects. Since a pronominal subject more often represents old information than a name subject does, the predicate for the pronominal subject more often represents new information and, by the same reasoning as above, tends to be longer. Comparing pronominal subjects to null subjects, we know that there are cases where pronouns represent new information. For example, they can serve to disambiguate

3.4 The Output Omission Model

The major property that distinguishes processing accounts (Bloom 1990, Valian 1991) from the series of accounts starting with Hyams 1983 is that the former do not accept that there is a stage where null subject sentences are grammatical for the English-speaking child. Nonetheless, children produce many sentences that lack subjects; and this fact must be explained. Moreover, we know that the subjects are there (are actually computed at some level) since, as L. Bloom (1970), Hyams (1986), and others have pointed out, the missing subjects have definite reference for the child.

Proponents of the processing-overload idea claim that an obligatory constituent is omitted because of "processing" constraints. However, no explanation is given for the omission process. Such an effect is not related to any theory or research on linguistic processing of which we are aware. Adult speakers do not "omit" constituents because of "processing" constraints. Apparently, then, this is a phenomenon unique to children, a phenomenon to which little actual research has been devoted. Hence, it is very difficult to discern its properties.

Since Bloom and other proponents do not provide a formal model, they make few detailed predictions about what such an "omission" stage should look like. So we will attempt the work of formalization here. Although we will try to be faithful to the ideas that underlie the explanation of null subjects, their lack of explicitness might lead to some ambiguity. We believe, however, that we have captured the ideas as well as they can be captured.

We will call the model the *Output Omission Model* (OOM). "Omission" because it is a model of omission of a constituent, in particular of a subject, in speech. "Output" because clearly the constituent must be computed at some level, and thus it is only in "output" that the constituent is omitted. For ease of exposition, we will sometimes refer to the model simply as the omission model.

Recall that here we are concerned only with the statistical properties of child speech as predicted by the omission model. We are not reviewing explanatory linguistic properties, although (in contrast to the grammatical model) the omission model says nothing about them.

Before we formalize the omission model (section 3.5), we will try to show intuitively what distinguishes it from the grammatical model, with respect to statistical properties. We have already discussed the subject-object asymmetry, which follows immediately

---

antecedence or to provide emphasis or contrast. This is not new information in the referential sense. Nevertheless, these pronominal subjects more often have predicates with old information than do null subject sentences in which the null subject always represents old information. Again, this results in statistically longer predicates for null subject sentences than for pronominal subject sentences.

Bloom (1990) considers the possibility that the VP length effect he obtains is a pragmatic one, but rejects this hypothesis. According to Bloom, the pragmatic hypothesis predicts that there should be no difference in VP length in sentences with pronouns (1st and 2nd person) versus sentences with lexical NP subjects because they all have unambiguous reference. The account we are suggesting, however, is not related to ambiguity of reference.

from the grammatical model, but for which the omission model needs empirically unjustifiable assumptions, so we will concentrate here only on statistical properties of the subject position.

The grammatical model sees a strong continuity between the use of null subjects by the young child and the use of pronouns by the older child and adult. In general, the young child uses null subjects in many (though not all) cases in which an older child or adult uses a pronoun. This is because, according to the grammatical model, the young child has a null subject language, and the use of null subjects in contexts in which a non–null subject language uses pronouns is simply a property of null subject languages.

The OOM, on the other hand, makes no such connection between the young child's use of phonetically empty subjects and the older child's use of pronouns. The model simply claims that a child will sometimes omit a subject, whether it be lexical or pronominal, when "processing load" is too great, in particular, when the VP is too long. Lexical subjects will be omitted at lower processing loads (VP lengths) than are pronouns, because the OOM stipulates (presumably, on phonetic grounds) that pronouns themselves impose less processing load than do lexical subjects.

Of course, under the OOM we still cannot predict the relative frequencies of null, pronominal, and lexical subjects, because other processes are at work besides omission. In particular, lexical or pronominal subjects will be selected, according to an array of factors, before the omission process even begins to apply. We can make use of one property of the model, however. The processing load asymmetry is localized to the omission process. Selection proceeds according to pragmatic, semantic, and situational factors, and then processing load determines whether the subject is omitted.

Suppose, therefore, that we observe a child in a state in which he is no longer omitting subjects, say, at a somewhat older age. We can use data from the child at that age to estimate the relative proportion of lexical and pronominal subjects. Suppose we observe that the child uses about two-thirds pronominal and one-third lexical subjects. Since the child uses no null subjects, we know that no omission has occurred. Therefore, all the subjects were selected as they appear, and we know that this child has selected two-thirds pronominal and one-third lexical subjects.

Consider table 4, which shows the proportions of lexical and pronominal subjects used by Adam and Eve during the developmental period under discussion. We see that, in fact, this 1:2 proportion of lexical to pronominal subjects is true for Adam 30;[30] so we know that the relative proportion of selected lexical subjects is about one-third. It is important to remember that this is the proportion of "selected" lexical subjects, not necessarily the proportion of "observed" lexical subjects. These two proportions only have to be the same when the child is not omitting any subjects, that is, after the null subject stage. At earlier stages the omission process is at work. In general, we expect pronominal and lexical subjects to be omitted at differential rates (i.e., it takes less of

[30] Eve at a late stage has more than 80% pronominal subjects. We will use Adam here for illustration.

Table 4
Proportions of lexical and pronominal subjects and derived bounds on L. (p_L = proportion of lexical subjects; p_P = proportion of pronominal subjects)

Sample	p_L	p_P	$min(L) = p_L/(p_L+p_P)$	$max(L) = 1 - p_P$	
ADAM					
06	.33	.11	.75	.89	*
08	.23	.20	.53	.80	*
10	.35	.20	.64	.80	*
12	.14	.20	.41	.80	*
14	.15	.15	.50	.85	*
16	.12	.52	.19	.48	
18	.16	.60	.21	.40	
20	.11	.77	.13	.23	*
30	.30	.67	.31	.33	
EVE					
02	.11	.29	.28	.71	*
04	.12	.37	.24	.63	*
06	.57	.14	.80	.86	*
08	.47	.26	.64	.74	*
10	.31	.37	.46	.63	*
12	.21	.68	.24	.32	*
14	.13	.74	.15	.26	
16	.23	.70	.25	.30	*
20	.11	.82	.12	.18	

* indicates the bounds for L that are impossible if L is estimated from the late stage (Adam 30, Eve 20)

a processing load to cause a lexical subject omission than a pronominal subject omission), so that the observed proportions of lexical and pronominal subjects will differ from the proportions of selected lexical and pronominal subjects.

We know, then, that Adam selects pronominal over lexical subjects by a factor of about two to one. We also know that the probability of omitting a lexical subject is greater than the probability of omitting a pronominal subject. Therefore, since observed subjects at any age are a result of two processes—selection and omission—we conclude that at ages when omission occurs (i.e., at the ages when subjects may be empty), the omission process must increase the number of pronominal subjects relative to the number of lexical subjects. That is, at younger ages the ratio of pronominal to lexical subjects must be *greater* than the ratio of selection (i.e., greater than two to one).

A look at the empirical results in table 4, however, tells quite a different story. Adam 06, for example, shows a 3:1 ratio in favor of lexical subjects, differing from the 1:2 ratio by a factor of at least 6. Similarly, Adam 08, 10, 12, and 14 all show impossible relationships according to the prediction just made. What is going wrong? Essentially, the older Adam is using more pronominal than lexical subjects, something we might expect. Since the subject often is at least to some extent "old" information, we expect quite a few pronouns in subject position. The younger Adam, however, uses very few pronouns; rather, he uses far too many lexical subjects.

The question is, Why the (relative) scarcity of pronouns in the younger Adam's speech? The answer follows directly from the grammatical analysis of the null subject stage. According to this theory, phonetically null subjects are not "omissions" of other subjects. Rather, they are licensed grammatically as empty subjects, in ways analogous to adult null subject languages. Basically, null subjects in these languages are used in most of the contexts where non–null subject languages use pronouns. We therefore expect a trade-off between null subjects and pronouns if Adam is developing from having a null subject language to having a non–null subject language. Roughly, we would expect the sum of the proportion of the null and pronominal subjects to be constant (up to the varying effects due to different situational contexts at each recording period). This is equivalent to saying that the proportion of lexical subjects should be roughly constant over the recordings at different ages. More precisely, there should be no trend of the proportion of lexical subjects decreasing or increasing from Adam 06 to Adam 30. And, in fact, this is exactly what we find. Adam 06 has 33% lexical subjects and Adam 30 has 30% lexical subjects. Eve 02 has 11% lexical subjects and Eve 20 has 11% lexical subjects.[31]

Here is another way to look at the problem. From the first to the last transcript the proportions of lexical subjects are about the same, and this is true for both Adam (.33 to .30) and Eve (.11 to .11). The proportions of pronouns, however, show a dramatic shift, for both Adam (.11 to .67) and Eve (.29 to .82).[32] Thus, the overall pattern of change from the null subject to the non–null subject stage is a dramatic increase in the number of pronominal subjects with a (roughly) steady number of lexical subjects. This is exactly what we would expect under the grammatical model, since null subjects trade off with pronouns under this theory. But it is vastly at odds with what we would expect under the OOM.

[31] Of course, the proportions vary over the transcripts from the first to the last. But we would expect that, since proportions of lexical subjects vary with the demands of the contextual situation. The important point is that there is no steady increase in the proportion of lexical subjects. And this is true for both Adam and Eve.

[32] Note that the last transcripts of Adam and Eve do not necessarily show them completely out of the null subject stage. This is especially true for Eve, since Eve 20 still has 7% null subjects. Thus, if we had later transcripts, the trends might be even more dramatic.

3.5 A Formalization and Analysis of the Output Omission Model

Having presented the intuitive basis of the OOM, we next derive equations that allow us to calculate explicit values of observed proportions. The reader who has understood the statistical inadequacies of the OOM given the verbal arguments of the previous section and who does not wish to pursue all the technical details may wish to skip this section, which formalizes and sharpens the argument. Section 3.6, however, presents new material on pronominal objects that has not been discussed previously.

The omission model specifies that the child always selects a lexical or pronominal subject (this is the grammatical process), but then may optionally omit it. First, therefore, we need to specify the probabilities with which the lexical or pronominal subject is selected. We do this in (7a). In (7b) we express the probabilities for deleting the different kinds of subjects.

(7) *The Output Omission Model*

a. *The Selection Process*
Select a lexical subject with probability L, or
Select a pronoun subject with probability $1 - L$.

b. *The Deletion Process*
If the subject is lexical, delete it with probability d_L.
If the subject is pronominal, delete it with probability d_P.

(7a) expresses the assumption that all (grammatically represented) subjects are either lexical or pronominal. There are no grammatical null subjects. This is the basic, crucial assumption of the OOM, and it is what distinguishes it from various versions of the grammatical theory. (7b) follows from the assumption that lexical subjects and pronoun subjects are deleted differentially because they induce different processing loads. Thus, there is one probability for omitting lexical subjects and a separate probability for omitting pronominal subjects.

In fact, the omission model predicts the VP length effect by assuming that lexical subjects have a higher processing cost than pronominal subjects. Therefore, according to the model, since deletion is the result of exceeding processing capacity, we can conclude that lexical subjects are deleted more frequently than pronominal subjects. This assumption is expressed in (8).

(8) $d_L > d_P$

We need to calculate the proportion of observed subjects of each kind. We will use the notation shown in (9).

(9) a. p_L = proportion of lexical subjects
b. p_P = proportion of pronoun subjects
c. $p_N = 1 - p_L - p_P$ = proportion of null subjects

The equation in (9c) follows because the three kinds of subjects (lexical, pronominal, and null) are exhaustive. Thus, their proportions must add up to 1.

We can now calculate the probabilities of different kinds of subjects, given the omission model (7). These calculations produce the equations in (10).[33]

(10) a. $p_L = L(1 - d_L)$
b. $p_P = (1 - L)(1 - d_P)$
c. $1 - p_L - p_P = Ld_L + (1 - L)d_P$

The crucial assumption in the OOM that makes these calculations possible is that there are two processes—selection and omission—that operate independently. To find the probability of two independent events, one simply multiplies their individual probabilities. Thus, consider (10a). The probability of observing a lexical subject is the probability of (a) selecting a lexical subject and (b) not deleting the lexical subject. The probability of (a) is L and the probability of (b) is $(1 - d_L)$. Multiplying these two probabilities yields (10a).

Similarly, the probability of observing a pronoun subject (10b) is the probability of (a) selecting a pronoun subject and (b) not deleting this pronoun subject. The probability of (a) is $(1 - L)$ and the probability of (b) is $(1 - d_P)$. Multiplying these two probabilities yields (10b).

(10c) represents the probability of observing a null subject. A null subject can be obtained in one of two ways: (a) Select a lexical subject (probability L) and then delete the lexical subject (probability d_L), with joint probability Ld_L, or (b) select a pronoun subject (probability $1 - L$) and then delete the pronoun subject (probability d_P), with joint probability $(1 - L)d_P$. The probability of a disjunction of two events is obtained by adding their probabilities. (10c) follows.

Equation (10c) is really not independent of equations (10a–b). Thus, (10) actually yields two equations in three unknowns: L, d_L, and d_P. We assume that p_L and p_P are estimated from the data (as the proportion of lexical and pronoun subjects, respectively).

However, we can escape this indeterminancy problem by taking advantage of the fact that the available data reach an age where the child is out of or almost out of the null subject stage. Consider Adam 30 in table 4. At that stage $p_L = .30$ and $p_P = .67$. Therefore, $(p_L + p_P)$, the proportion of nonnull subjects, is .97. We can therefore approximate Adam's L by the proportion of lexical subjects at Adam 30; that is, we can take $L = .30/.97 = .31$.

If we substitute $L = .31$ into (10) and solve, using the observed proportions of p_L and p_P, we find that the equations do not solve appropriately. That is, many of the

[33] (10) is actually a calculation of the probabilities of particular events; for example, (10a) is the probability of observing a lexical subject. We have identified the probability of an event with the proportion of observed events of that type, because it is clear that the proportion of observed events is the appropriate estimate in this case of the probability of the event. Thus, for example, the proportion of observed pronouns is the best estimate of the probability of a pronoun appearing in subject position.

solutions for d_L and d_P turn out to be negative numbers, and thus not probabilities. This shows that the OOM does not fit the statistical data. The same is true for the data from Eve, using an estimate of L from Eve 20.

Let us try to show the problem in a clearer way. We will derive bounds on L and show that L does not fall within these bounds. First, we can calculate (11) from the model (7), or (10).

(11) $L < 1 - p_P$

We will not show an algebraic derivation, but the truth of (11) can be seen very simply. Namely, $(1 - p_P)$ is the sum of the proportions of lexical and null subjects. If a lexical subject is selected (which happens with probability L), it will be deleted or not deleted; that is, it will end up observed as either a lexical subject or a null subject. So the probability of selecting a lexical subject cannot be higher than the total proportion of lexical or null subjects. In fact, L will become as high as $(1 - p_P)$ only if the probability of deleting a pronominal subject is zero, and thus, all null subjects are derived from lexical subjects.

In addition to the upper bound (11), we can also derive a lower bound for L, namely, (12), which is calculated by substituting (8) into (10) and solving.

(12) $L > p_L / (p_L + p_P)$

The intuitive basis for (12) may be seen from the following. The right-hand side is the relative proportion of lexical subjects out of all nonnull (i.e., lexical or pronominal) subjects. Since (8) says that lexical subjects are omitted more frequently than pronominal subjects, in order to get this particular observed proportion of lexical subjects, a greater proportion of lexical subjects would have had to be selected. For example, suppose the right-hand side of (12) is .30. That means that .30 of the nonnull subjects are lexical. Suppose, however, that L is less than .30, say, .20. That means that lexical subjects are selected 20% of the time. But then, when omission occurs, more lexical subjects are omitted than pronominal subjects. Therefore, the observed proportion of lexical to non-null subjects will have to be less than 20%.

In the fourth column of table 4 we have listed *min(L)*, the minimum value of L, determined by the right-hand side of (12). In the fifth column we have listed *max(L)*, the maximum value of L, determined by the right-hand side of (11). For example, consider Adam 06. Given the observed proportions of lexical (.33) and pronominal (.11) subjects, we can calculate from (12) that L must be greater than .75 and from (11) that L must be less than .89. Similarly for the other transcripts. As can be seen, this puts rather tight bounds on L. In particular, at Adam 06 it means that lexical subjects must be selected at least 75% of the time, which is at variance with what we have seen for Adam at the non–null subject stage (Adam 30), when L can be estimated directly (at about .31).

In particular, suppose we consider the bounds on L for Adam from Adam 30 and the bounds on L for Eve from Eve 20. Adam's bounds on L are between .31 and .33.

Eve's bounds are between .12 and .18. These are consistent with the observation that subjects are very often pronominal, rather than lexical. Now we can ask, Which of the earlier transcripts show bounds on L that could be compatible with the bounds from the later, non–null subject stage? We have put an asterisk next to each transcript that is not compatible with the bounds estimated from the last transcript. For example, there is an asterisk next to Adam 06 because there is no point in the bounds for Adam 06 (.75 to .89) that is also in the bounds for Adam 30 (.31 to .33). There is an asterisk next to Eve 06, because there is no point in the bounds for Eve 06 (.80 to .86) that is also in the bounds for Eve 20 (.12 to .18). There is no asterisk next to Eve 14 because there are points in the bounds for Eve 14 (.15 to .26) that are also in the bounds for Eve 20 (.12 to .18), namely, the interval from .15 to .18.

As can be seen from table 4, out of eight transcripts for Adam (not counting Adam 30, since that is the transcript from which we estimate), there are six transcripts whose boundaries on L are not compatible with those estimated from Adam 30. Similarly, out of eight transcripts for Eve (not counting Eve 20), there are seven transcripts whose boundaries on L are not compatible with those estimated from Eve 20. In fact, for both Adam and Eve, many of the numbers are very far off. Even though $max(L)$ for Adam is estimated as .33, the numbers for Adam 06 to Adam 10 are .75, .53, and .64, respectively. Although $max(L)$ for Eve is .18, the minimum values for the transcripts range as high as .80.

Thus, we can only conclude that the statistical properties of the OOM are widely at variance with the data. The basic problem is clear. The non–null subject stage has a large majority of pronouns in subject position. Basically, these pronouns appear where earlier there were null subjects. The OOM does not capture this process since it has no concept of (grammatical) null subject. It posits only omitted subjects, which can derive either from lexical or from pronominal subjects.

3.6 Pronoun Difficulty Factor

An attempt might be made to save the OOM with respect to the statistical data by adding a large number of parameters to the model, basically by letting L (the proportion of selected lexical subjects) vary with age. Such a proposal is theoretically undesirable in that it would add large numbers of degrees of freedom to the model, thereby robbing it of predictive power. However, we can actually test the empirical validity of the claim that the younger the child, the more likely she is to select a lexical subject. Suppose it were true that young children select a much larger ratio of lexical to pronominal subjects than do older children. Let's say this has to do with a *pronoun difficulty factor*—PDF. Though we will not attempt to specify PDF, keep in mind that it is a factor that impedes the *selection* of pronouns according to the OOM, not the *output* of pronouns. It has nothing to do with position in the sentence, or similar factors; it simply creates a relative difficulty for pronoun selection compared to lexical NP selection. Therefore, PDF will also inhibit the selection of pronouns for grammatical objects. Since PDF is higher for young children (based on the results for the selection of subjects), the prediction is that young children will also select a much larger ratio of lexical objects to pronominal objects than will older children. In short, the prediction is that the same growth of selected pronouns relative to lexical NPs will occur for objects as for subjects.

In fact, since relatively few objects are omitted, the effects of the (pronominal and lexical) deletion parameters will be quite a bit smaller. Therefore, we will be able to derive even tighter bounds on L—the probability of selecting a lexical NP—for objects than we did for subjects. To see this, consider the limiting case where no objects are deleted. In this case we would be able to estimate L for objects exactly as the proportion of lexical NPs.

Table 5 shows the proportions of lexical and pronominal objects for Adam and Eve, and also the minimal and maximal values of L (the probability of selecting a lexical object), according to the OOM. The calculations are done exactly as for subjects in table 4. For the OOM, the differences between subject and object production are in the deletion

Table 5
Proportions of lexical and pronominal objects and derived bounds on L. (p_L = proportion of lexical objects; p_P = proportion of pronominal objects)

Sample	p_L	p_P	$min(L) = p_L/(p_L+p_P)$	$max(L) = 1 - p_P$
ADAM				
06	.50	.44	.53	.56
08	.68	.29	.70	.71
10	.68	.21	.76	.79
12	.61	.31	.66	.69
14	.44	.43	.51	.57
16	.33	.54	.38	.46
18	.43	.43	.50	.57
20	.74	.20	.79	.80
EVE				
02	.58	.21	.73	.79
04	.62	.27	.70	.73
06	.70	.19	.79	.81
08	.65	.21	.76	.79
10	.75	.18	.81	.82
12	.62	.32	.66	.68
14	.89	.07	.93	.93
16	.80	.19	.81	.81

parameters (d_L and d_P), which will be different for subjects and objects. There may also (independently of the OOM) be a difference in the probability of selecting a lexical subject versus a lexical object, and similarly for pronouns. However, the OOM must predict that PDF will cause the older child to select relatively fewer lexical objects than does the younger child.

Is this prediction borne out? Strikingly, no. There are major differences between the patterns for L in table 4 (subjects) and table 5 (objects). First consider Adam. Although we do not know the exact values of the deletion parameters, we were still able to derive bounds on L. For Adam 06's subjects (table 4) we know that L must be at least .75 and for Adam 30 L may be no greater than .33. In fact, the minimum values of L for the three youngest transcripts (06, 08, 10) are all greater than the maximum values of L for the four oldest transcripts (16, 18, 20, 30). As discussed earlier, this is a striking case of decrease in L.

Now let us look at table 5, for objects. First, the small number of null objects means that we can derive even tighter bounds on L. Is there a pattern of L decreasing with age? Quite strongly, no. In fact, the youngest transcript, Adam 06, shows a fairly low, not high, L (between .53 and .56). The oldest transcript, Adam 20, shows the highest interval of all (.79 to .80). Whatever one thinks about the pattern of L values, it is quite clear that there is not a strong decrease with age, as there is for subjects.[34]

Now consider Eve. Again, for subjects (table 4), we see a decrease in L values with age. Because of the large number of null subjects and the fact that Eve often selects more pronominal than lexical subjects, the predicted intervals are fairly wide. However, the oldest transcript (Eve 20) shows a maximal value of L (.18) that is smaller than all but one of the minimal values. In fact, the intervals are such that it could easily be (almost must be) that the last four values (Eve 12–20) are each smaller than each of the first five values (Eve 02–10). Of the first five transcripts, it would only be possible for Eve 02 and Eve 04 to have lower values than any of the last four transcripts; and for this effect to occur, extreme values of the deletion parameters would have to be chosen, a quite unlikely result. In summary, it is quite clear that for Eve also, L decreases with age, according to the OOM.

Is it also true that L decreases with age for Eve's objects? Again, no.[35] The largest value of L is found at Eve 14 (.93, an exact value since there were fewer than 1% null objects at this stage). The oldest age (Eve 16) essentially ties for second largest value of L. Basically, these data do not show much of a trend—if anything, however, L goes up with age, not down. Again, the declining pattern that occurs with subjects is not repeated.

[34] We could calculate the relation of L value to age by computing an r value between L and age. However, we have not done so for two reasons. First, it is not clear what L value to use, since the OOM predicts only an interval, not a particular value, and the intervals are even wider for subjects. Second, for each child, there are only eight values to enter into the calculation. However, the trend or lack of trends is quite clear.

[35] We have omitted Eve 20 from table 5, because there was a fairly small sample in this transcript, and thus not many objects. The proportions were p_L = .55, p_P = .43, $min(L)$ = .56, and $max(L)$ = .44.

A striking property of both Adam's and Eve's objects, to be seen in table 5, is the roughly steady proportion of object pronouns as age increases. We pointed out earlier that both Adam and Eve show a large growth in the proportion of subject pronouns with age. This is what the grammatical model predicts, since pronouns replace earlier null subjects. However, there is no reason for the grammatical model to make such a prediction about objects, and this is consistent with the results. Nongrammatical models like the OOM must explain the growth in subject pronouns as a result of PDF, which then should also apply to objects. The lack of growth in object pronouns is quite clear empirical evidence against the OOM and its necessary assumption that young children select proportionately more lexical NPs, a claim we have already argued would be suspect on methodological and conceptual grounds.

It is difficult to see how the OOM can be saved from this further empirical argument against it. One could try to specify yet more parameters of the model that would result in the young child selecting proportionately more lexical subjects than the older child, but not proportionately more lexical objects. In other words, not only would the selectional parameters for subjects and objects differ, but their growth patterns would differ as well. This would amount to saying that PDF held for subjects but not for objects, and one could try to justify this assumption. However, in addition to the methodological problem of the explosion of parameters in this model, it is impossible to assume that there is no PDF for objects—such an assumption would be completely against the spirit, the theoretical underpinnings, of the OOM. Recall that L is the selectional parameter, not a "processing" parameter. Children (like adults) will select lexical or pronominal subjects for a number of pragmatic and semantic reasons. Processing considerations that distinguish pronouns from lexical NPs are localized in the deletion parameters d_L and d_P. If there is a processing consideration that affects the selection of a subject pronoun at one age, then this processing consideration must affect the selection of an object pronoun at that age.

The grammatical model, on the other hand, finds the results on objects quite compatible. There are increasing numbers of subject pronouns as children grow older because they replace null subjects. Lexical subjects are roughly constant. Pronominal and lexical objects are roughly constant because there is no grammatical null object process. The pattern of the data is what one would expect from the grammatical model.

In sum, (a) it makes little sense to assume that the youngest children select lexical NPs far out of proportion to pronouns, compared to older children and adults, (b) methodologically, the model becomes quite unconstrained and untestable if one makes such an assumption, and (c) the evidence from object position argues against such an assumption. We must conclude that there is no reasonable way to bring the OOM into accord with the empirical results.

We have shown that the OOM fails with respect to the statistical properties of child language production, properties it was set up to capture (i.e., the VP length effect). Moreover, it fails in a very perspicuous way. That is, it fails because it does not capture the clear fact that there is a trade-off in early child speech: pronouns replace null subjects as children get older. In an attempt to save the OOM, one might be forced to adopt a model with a large number of degrees of freedom, and hence with little or no predictive capacity. The grammatical model, in contrast, captures the trade-off in a natural way.[36]

4 Conclusion

It is a trivial observation that children are limited in their productive abilities. We have shown, however, that this observation does not constitute an explanation for the missing subject phenomenon. The obvious question (posed by Feldman, Goldin-Meadow, and Gleitman (1978)) is, Why should it be the subject (actor, in their terms) that is selected for omission when there is competition for space? In other words, why do we find the subject-object asymmetry discussed earlier?

We have shown that there is little theoretical or empirical motivation for the claim that the beginning of the sentence is more difficult to process than the end, and thus that the processing account provides no explanation for the striking asymmetry in argument omission. Moreover, the asymmetry cannot be explained solely on the basis of Informativeness considerations. In addition, we showed that in contrast to the various grammatical accounts of null subjects, the processing and Informativeness analyses offer no principled explanation for the fact that null subjects typically occur alongside a range of other theoretically related grammatical properties. Finally, we presented a formal processing model of argument omission (the OOM) based on the general assumptions of the processing approach and showed that such a model is simply incompatible with the statistical data.

We certainly do not exclude the possibility that performance factors may be involved in the null subject phenomenon in child language. Studies of adult null subject languages have shown that a number of pragmatic constraints operate in this domain. Indeed, we suggested that the variation in VP length as a function of subject type may be due to

[36] Bloom, Miller, and Hood (1975) propose a variable rules model in the manner of Labov (1969) to account for argument omission in child language. The model they propose specifies the probability of deletion as determined by a number of conditioning factors: grammatical (see section 3), discourse support (whether there is a preceding related sentence), and lexical familiarity.

This model is difficult to evaluate in the present context since it is unclear whether it is intended as a competence model or a performance (production) model. Bloom, Miller, and Hood claim that the omission of constituents (reduction, in their terms) is a "grammatical process" (p. 46). However, it is clear that at least two of the three conditioning factors (discourse support and lexical novelty) are not grammatical. Moreover, the notion of "grammatical complexity," as used in this model, is more of a processing notion than a grammatical one. Indeed, Bloom, Miller, and Hood note that "the variable rules in the model of child grammar suggested here are, in some essential sense, related to the development of memory processes" (p. 50).

Thus, the variable rules model proposed by Bloom, Miller, and Hood (and by Labov, for that matter) is not, in fact, a model of the child's grammatical knowledge, at least not under any standard construal of the notion of competence. As a processing or performance model, however, it fails in much the same way as the proposals discussed earlier; that is, if we grant that there are discourse, lexical, and complexity factors that affect production, why is it overwhelmingly the subject that is omitted under this extragrammatical pressure? And why is there a trade-off between null subjects and pronouns?

pragmatic factors (footnote 29), and we also proposed an explanation for the null subject phenomenon in terms of an interaction of structural and pragmatic factors (section 2.3). Moreover, since children are clearly limited in their productive abilities, it is possible that this provides a functional explanation for why they initially assume a positive value for the null subject parameter (Weinberg 1987, Rizzi 1986).[37] However, the performance models discussed in this article, the OOM in particular, fail to provide an adequate account of the range of properties associated with null subjects in child language; thus, whatever the contributions of Informativeness and processing constraints to the null subject phenomenon, they do not vitiate the need for a principled grammatical account. It seems to us that the basic insight that relates null subjects in child speech to null subjects in adult speech is correct and that the grammatical approach to this problem is the only one that has achieved a level of descriptive and explanatory adequacy.

Appendix: The Output Omission Model and Yngve's (1960) Production Model

Bloom (1990) suggests that Yngve's (1960) model of sentence production supports the claim that the beginning of the sentence is harder to process than the end. Yngve's model assumes that, in production, categorial nodes are produced top-down and left-to-right. Furthermore, memory load (a computational burden) at any node in a partially generated phrase marker is proportional to the number of unexpanded nodes dominating that node. As Fodor, Bever, and Garrett (1974) point out, the main empirical consequence of this model is that left branches should impose particular memory demands on speakers. Bloom suggests that this model also predicts serial (beginning vs. end of sentence) effects. However, reviewing Yngve's model, we find no prediction that a higher memory load is imposed on the sentence during production of the first word or constituent.[38]

[37] Weinberg (1987) has suggested that processing limitations may provide a functional explanation for why null subjects represent the initial setting of the null subject parameter. She suggests that because children have a limited processing capacity, they avail themselves of a grammatical option that imposes a lesser processing load.

Rizzi (1986) makes a suggestion along similar lines. He notes that although [− null subject] may constitute the unmarked case (according to his grammatical analysis), children may actually opt for the marked [+ null subject] setting because of a limited ability to process the input data. As their processors mature, providing them with more complete access to data, English-speaking children switch to the unmarked [− null subject] setting. Both Weinberg and Rizzi assume that the child's grammar differs from the adult's in permitting null subjects, which is different from the position of Bloom (1990) and Valian (1991), who maintain that the processing limitations hypothesis vitiates the need for a grammatical analysis.

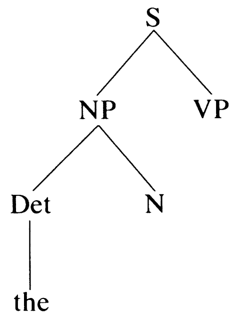
Mazuka et al. (1986) also propose that subject omission results from the interaction of grammatical properties of a particular language (i.e., branching direction) and a length/complexity constraint. According to their analysis, children are limited in their productive abilities. In a right-branching language such as English subjects are omitted because complexity builds in a rightward direction, and hence it is the elements on the left that are reduced or omitted.

There are a number of problems with Mazuka et al.'s account (see Hyams 1987). For example, it predicts that in a left-branching language such as Japanese children should show a preference for omitting verbs, which seems not to be the case.

[38] We are concerned here only with whether Yngve's model would predict that the first word or constituent of a sentence should be omitted, or at least, has a higher memory load. We do not discuss whether Yngve's model is an appropriate model for production. In this regard, we should note that, contrary to what Bloom

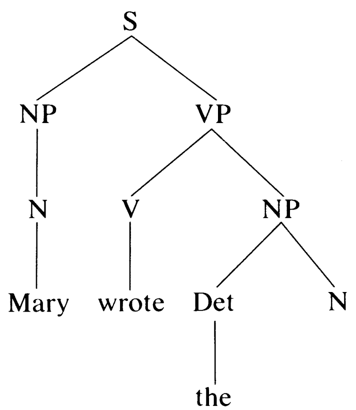
To see how Yngve's model works, consider an example provided by Fodor, Bever, and Garrett (1974:407–408). Slightly updating their notation, suppose that the partial phrase marker in (13) has been generated. That is, the speaker has expanded S into NP VP, NP into Det and N, and Det into *the*. This is in accordance with the top-down and left-to-right nature of the production process.

(13)



Fodor, Bever, and Garrett write, "According to Yngve's model, there are two stored nodes at the point in the production of this tree where the speaker says *the*." In other words, the necessity of keeping these two unexpanded nodes in memory imposes a certain memory load on the speaker. It can be seen that the memory load on a lexical item is proportional to the number of left branches dominating that item. Suppose, in fact, that a child omitted words at points of high memory load, that is, points at which there were a large number of unexpanded nodes. Compare the situation in (13), where there are two unexpanded nodes, with a partially generated phrase marker in which the N in the subject had been produced, the VP had been expanded as V NP, and the V had been expanded with a lexical verb, as in (14).

(14)



(1990) asserts, there has been no strong empirical support for Yngve's model over the past 30 years. Chomsky (1965:197–198, n. 8) finds Yngve's model "implausible" and "without empirical support." Fodor, Bever, and Garrett (1974:406–419) conclude that "the experimental literature still lacks an entirely convincing demonstration that left-branching structures are harder to produce than their right-branching counterparts." And in Levelt's (1989) comprehensive work on production, Yngve's model is not even mentioned.

In (14) there is only one left branch dominating *Mary* and one left branch dominating *wrote*. Therefore, we would expect omission of *wrote* as often as omission of *Mary*. Of course, this is not what we find.

To see this in more detail, note that if we adopt Yngve's model in conjunction with the OOM, we are forced to make the assumptions in (15).

(15) a. The more left branches dominating a lexical item, the more likely the item is to be omitted.
b. For a given memory load, lexical NPs are more likely to be omitted than pronouns.

Now, why should lexical subjects be omitted? Suppose that the subject is just one word—say, a proper noun like a name, or a common noun without a determiner, as is often found in child speech at the age in question. Then there is only one left branch dominating the subject. Only S must be kept in memory.[39] In other words, for lexical subjects that contain only one word, the memory load is no more than for verbs. Why are these lexical subjects omitted so much more often than verbs?[40]

It might be claimed that what is being omitted is not lexical subjects with only one word, but more complex subjects—say, those composed of a Determiner and a Noun. But suppose that NP is expanded into Det N, as in (13). Then the high memory load (15b) occurs when the lexical item filling Det is ready to be pronounced. So *the* should be omitted. But then the lexical item filling the Noun can be pronounced, with at most a memory load of 1 affecting it (the unexpanded S). (See footnote 38 on why in fact there might be a zero memory load here.)

We have deduced that the memory load on a subject Noun, whether a single word (N) subject, or a subject Noun preceded by a Determiner, is not more than 1, the same memory load as on a verb. Yet verbs are omitted far less frequently than subjects.

What about Yngve's model makes subjects as a whole difficult? The left branch prediction would pick out determiners in subject position at best, but not the subject overall. So there seems to be no way, even making the assumptions in (15), to predict that a whole subject will be omitted. Even more strikingly, recall that the subject *must* be computed for the correct interpretation of the sentence. That is, the subject must be generated in a phrase marker like the ones we have been exemplifying. Moreover, if children could compute a VP without a subject, they would be computing an ungrammatical utterance. This is exactly what the OOM is trying to avoid; that is, the OOM holds that the child's grammatical system is like the adult's. In fact, the situation would be worse than the null subject hypothesis in this regard, since it is not clear that S dominating a VP and no subject is at all possible in UG.

[39] In fact, when the subject noun is produced, this completes the NP and the generator moves up to S and begins to work on VP. Depending on assumptions, we might say that when the subject is produced, there is actually *no* memory load on it.

[40] Bloom, Miller, and Hood (1975) report the relative frequencies of various constituents (e.g., agent, verb, object) in the utterances of the four children they studied. The rate of verb omission is vastly lower than that of subject omission. See Bloom, Miller, and Hood 1975 for further discussion.

So Yngve's model really can have nothing to do with the OOM. Rather, the OOM assumes that pronunciation is difficult when there is a memory load complexity induced by the actual representational computation during production. This is rather odd, since the point of production cannot possibly be the point of computation of the subject. At the point where a child is trying to pronounce a subject (a fact also pointed out to us by Fernanda Ferreira), the computation of the surface structure has proceeded beyond the subject. Why should it be difficult to pronounce a subject while there is a lower computational load, due to the procedure currently generating a phrase beyond the subject?

The assumptions of the OOM are in fact at odds with Yngve's model. To see this, recall that the latter model, working top down and left to right, produces subjects before VPs. But recall also that according to the OOM, subjects are differentially omitted as a function of VP length. Therefore, according to the OOM, the VP must be selected before the subject. Thus, if the OOM is looked upon as a production model for a sentence, then it is completely incompatible with Yngve's model and thus the latter can hardly be used to justify Bloom's (1990) claim that subject position has a larger computational load than object position.

Of course, as we have pointed out, the proper way to look at the OOM is not as a production model for the structure of a sentence, but rather as a (phonetic) omission model, which omits subjects as a function of VP length and phonetic weight after the structure has already been computed. However, under this (correct) interpretation, the entire phrase marker has already been generated at the time the OOM applies. Thus, the computational load described by Yngve's model (e.g., the effect of left branches) is irrelevant at the point of constituent omission. Therefore, Yngve's model, or any other model of phrase structure generation, cannot be responsible for the differential amounts of subject and object omission.

To sum up, the OOM, in conjunction with Yngve's (1960) processing model, fails to account for the differential rate of subject and object omission—the central traditional statistical fact concerning null subjects. Moreover, it fails to account for the trade-off between null subjects and pronouns discussed earlier. Finally, it raises—in a rather striking way—the question of why the phonetic weight of a subject should have anything to do with its computation.

References

Adone, Dany, and Maaike Verrips. 1989. The acquisition of empty categories in Mauritian Creole. Paper presented at the Boston University Conference on Language Development, October 1989.

Bellugi, Ursula. 1967. The acquisition of negation. Doctoral dissertation, Harvard University, Cambridge, Mass.

Bever, Thomas. 1970. The cognitive basis for linguistic structures. In *Cognition and the development of language,* ed. John Hayes, 279–362. New York: Wiley.

Bloom, Lois. 1970. *Language development: Form and function in emerging grammars*. Cambridge, Mass.: MIT Press.

Bloom, Lois, Patsy Lightbown, and Lois Hood. 1975. Structure and variation in child language. *Monographs of the Society for Research in Child Development* 40.2.

Bloom, Lois, Peggy Miller, and Lois Hood. 1975. Variation and reduction as aspects of competence in language development. In *The 1974 Minnesota Symposium on Child Psychology*, ed. A. Pick, 3–55. Minneapolis, Minn.: University of Minnesota Press.

Bloom, Paul. 1990. Subjectless sentences in child language. *Linguistic Inquiry* 21:491–504.

Borer, Hagit, and Kenneth Wexler. 1992. Biunique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory* 10:147–190.

Bowerman, Melissa. 1973. *Early syntactic development*. Cambridge: Cambridge University Press.

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, Mass.: Harvard University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Clahsen, Harald. 1986. Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics* 24:79–121.

Curtiss, Susan, and Paula Tallal. 1991. On the nature of the impairment in language-impaired children. In *New directions in research on child language disorders*, ed. Jon Miller. Boston: College Hill Press.

Diesing, Molly. 1988. Bare plural subjects and the state/individual contrast. In *Proceedings of the Tübingen Conference on Generics*, ed. Manfred Krifka. Natürlichsprachliche Institut für Systeme der Universität Tübingen.

Feldman, Heidi, Susan Goldin-Meadow, and Lila Gleitman. 1978. Beyond Herodotus: The creation of language by linguistically deprived deaf children. In *Action, symbol and gesture: The emergence of language*, ed. A. Lock. New York: Academic Press.

Ferreira, Fernanda. 1991. Effects of length and syntactic complexity on initiation times for prepaced utterances. *Journal of Memory and Language* 30:210–233.

Ferreira, Fernanda, and Frederick J. Morrison. 1990. Children's knowledge of syntactic constituents. Paper presented at the Annual Meeting of the Psychonomic Society, New Orleans.

Fodor, Jerry, Thomas Bever, and Merrill Garrett. 1974. *The psychology of language*. New York: McGraw-Hill.

Frazier, Lyn. 1985. Syntactic complexity. In *Natural language parsing*, ed. David Dowty, Lauri Karttunen, and Arnold Zwicky, 129–189. Cambridge: Cambridge University Press.

Gerken, LouAnn. 1991. The metrical basis for children's subjectless sentences. *Journal of Memory and Language* 30:431–451.

Gibson, Edward, and Kenneth Wexler. 1992. Parameter-setting, triggers and V2. Paper presented at the 15th GLOW Colloquium, Lisbon. Ms., MIT, Cambridge, Mass.

Greenfield, Patricia, and Joshua H. Smith. 1976. *The structure of communication in early language development*. New York: Academic Press.

Guasti, Maria Teresa. 1992. Finite and infinitive verbs in the acquisition of Italian. Ms., University of Geneva.

Guilfoyle, Eithne. 1984. The acquisition of tense and the emergence of lexical subjects in child grammars. In *McGill working papers in linguistics*. Linguistics Department, McGill University, Montreal, Quebec.

Guilfoyle, Eithne, and Máire Noonan. 1988. Functional categories and language acquisition. Paper presented at the Boston University Conference on Language Development.

Haan, Ger Jan de, and Kees Tuijnman. 1988. Missing subjects and objects in child grammar. In *Language development,* ed. Peter Jordens and Josine Lalleman, 101–122. Dordrecht: Foris.
Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15:531–574.
Hyams, Nina. 1983. Acquisition of parameterized grammars. Doctoral dissertation, CUNY, New York.
Hyams, Nina. 1986. *Language acquisition and the theory of parameters.* Dordrecht: Reidel.
Hyams, Nina. 1987. A reanalysis of null subjects in child language. In *Theoretical studies in language acquisition: Proceedings of the Berlin Workshop,* ed. Helen Goodluck, Thomas Roeper, and Jürgen Weissenborn. Hillsdale, N.J.: Lawrence Erlbaum Associates.
Hyams, Nina. 1991. The genesis of clausal structure. In *The acquisition of verb placement,* ed. Jürgen Meisel, 371–400. Dordrecht: Kluwer.
Hyams, Nina. 1992. V2, null arguments and C projections. To appear in *Language acquisition studies in generative grammar,* ed. Teun Hoekstra and Bonnie Schwartz. Amsterdam: John Benjamins.
Jaeggli, Osvaldo, and Nina Hyams. 1988. Morphological uniformity and the setting of the null subject parameter. In *Proceedings of NELS 18,* 238–253. GLSA, University of Massachusetts, Amherst.
Jaeggli, Osvaldo, and Ken Safir. 1989. The null subject parameter and parametric theory. In *The null subject parameter,* ed. Osvaldo Jaeggli and Ken Safir, 1–44. Dordrecht: Kluwer.
Kitagawa, Yoshihisa. 1986. Subjects in Japanese and English. Doctoral dissertation, University of Massachusetts, Amherst.
Koopman, Hilda, and Dominique Sportiche. 1991. The position of subjects. *Lingua* 85:211–258.
Kratzer, Angelika. 1989. Stage level and individual level predicates. In *University of Massachusetts papers on quantification,* ed. Emmon Bach, Angelika Kratzer, and Barbara Partee. NSF Grant Report.
Labov, William. 1969. Contraction, deletion and inherent variability of the English copula. *Language* 45:715–762.
Lebeaux, David. 1987. Comments on Hyams. In *Parameter setting,* ed. Thomas Roeper and Edwin Williams, 23–40. Dordrecht: Reidel.
Levelt, Willem. 1989. *Speaking: From intention to articulations.* Cambridge, Mass.: MIT Press.
Lillo-Martin, Diane. 1986. Effects of the acquisition of morphology on syntactic parameter setting. In *Proceedings of NELS 16,* 305–321. GLSA, University of Massachusetts, Amherst.
MacWhinney, Brian, and Catherine Snow. 1985. The Child Language Data Exchange System. *Journal of Child Language* 12:271–296.
Mazuka, Reiko, Barbara Lust, Tatsuko K. Wakayama, and Wendy Snyder. 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.
Meisel, Jürgen. 1987. Early differentiation of languages in bilingual children. In *Bilingualism across the lifespan: In health and pathology,* ed. Kenneth Hyltenstam and Lorraine Obler, 13–40. Cambridge: Cambridge University Press.
Newport, Elissa, Harry Gleitman, and Lila Gleitman. 1977. Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In *Talking to children: Language input and acquisition,* ed. Catherine Snow and Charles Ferguson, 109–149. Cambridge: Cambridge University Press.
Pierce, Amy. 1992. *Language acquisition and syntactic theory: A comparative analysis of French and English child grammars.* Dordrecht: Kluwer.

Pinker, Steven. 1984. *Language learnability and language development.* Cambridge, Mass.: Harvard University Press.

Poeppel, David, and Kenneth Wexler. 1991. Finiteness and V2 effects implicate the existence of functional categories and head movement in early German grammar. Paper presented at the Boston University Conference on Language Development, Boston.

Poeppel, David, and Kenneth Wexler. 1993. The full competence hypothesis of clause structure in child German. *Language* 69:1–33.

Radford, Andrew. 1990. *Syntactic theory and the acquisition of English syntax: The nature of early grammars of English.* Oxford: Blackwell.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro. Linguistic Inquiry* 17:507–557.

Rizzi, Luigi. 1992. Early null subjects and root null subjects. Ms., University of Geneva.

Schaeffer, Jeannette. 1990. The syntax of the subject in child language: Italian compared to Dutch. Master's thesis, University of Utrecht.

Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40:21–81.

Wang, Qi, Diane Lillo-Martin, Catherine T. Best, and Andrea Levitt. 1992. Null subjects and objects in the acquisition of Chinese. *Language Acquisition* 2:221–254.

Weinberg, Amy. 1987. Comments on Borer and Wexler. In *Parameter setting,* ed. Thomas Roeper and Edwin Williams, 173–188. Dordrecht: Reidel.

Weissenborn, Jürgen. 1990. Functional categories and verb movement. In *Spracherwerb und Grammatik,* ed. Monika Rothweiler. *Linguistische Berichte,* Sonderheft 3/1990.

Weverink, Meike. 1989. The subject in relation to inflection in child language. Master's thesis, University of Utrecht.

Wexler, Kenneth. 1991. Optional infinitives, head movement and the economy of derivations in child grammar. Occasional Paper 45, Center for Cognitive Science, MIT, Cambridge, Mass.

Wexler, Kenneth. 1993. Optional infinitives, head movement and the economy of derivations in child grammar. To appear in *Verb movement,* ed. Norbert Hornstein and David Lightfoot. Cambridge: Cambridge University Press.

Yngve, Victor. 1960. A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society* 104:444–466.

(Hyams)
Department of Linguistics
UCLA
Los Angeles, California 90024

(Wexler)
E10-020
MIT
Cambridge, Massachusetts 02139