Commentary: Null Subjects in Child Language and the Implications of Cross-Linguistic Variation

Nina Hyams
University of California at Los Angeles

A central point that has emerged from the chapters on null subjects is that although children universally omit subjects, there is also an interesting range of cross-linguistic variation in the young child’s knowledge and use of null and overt pronouns. Pierce (this volume, chapter 14) notes, for example, that subject pronouns in early French and early English differ in distribution and frequency. Valian (this volume, chapter 11) observes that although American and Italian children both use null subjects, Italian children have a much higher frequency of null subject use than English-speaking children at a comparable developmental stage. Lillo-Martin (this volume, chapter 13) points out that although both English- and Chinese-speaking children have null subjects, only Chinese-speaking children drop objects, as is permissible in the adult language (cf. Wang, Lillo-Martin, Best, & Levitt, 1992). Finally, Rizzi (this volume, chapter 10) claims, based on claims in Valian (1991) and Roeppe and Weissenborn (1990), that null subjects in early English are restricted to root contexts, whereas in null subject languages such as Italian, they are not so restricted. What these examples show is that child languages, like adult languages, show a complex interaction of universal and language-specific, that is, input-determined, properties.

In general, there are two perspectives from which to approach this kind of cross-linguistic variation. On the one hand, we might try to explain the variation we observe in child language the way we explain variation across adult languages, namely, as a reflex of differences and similarities in the underlying grammars. For ease of exposition, I refer to this as the grammatical approach. Pierce’s analysis of subject pronouns is an example of
such an approach. Pierce argues that subject pronouns in early French are inflectional heads, whereas in early English they are NPs. In this respect the early grammars mirror the adult grammars of these two languages, and the distributional differences in the acquisition data follow as a consequence of the different grammatical representations.

As noted earlier, there are also cross-linguistic differences in the child's use of null subjects. Although null subjects appear to be a universal property of early language, the structural conditions under which null arguments are licensed varies across child languages (as they do for adult languages). Thus, de Haan and Tuijnman (1988) and Poeppel and Wexler (1993) showed that children acquiring V-2 languages such as German and Dutch have topic-drop grammars; that is, null subjects (and objects) are licensed in Topic (i.e., [SPEC, CP]) position. This is in contrast to pro-drop languages such as Italian in which null subjects are licensed in [SPEC, [IP]]. Rizzi analyzes null subjects in early English as a kind of "diary drop," a phenomenon that also exists in adult English (Haegeman, 1990) and is distinct from both the pro-drop we find in Italian and the topic-drop of German and Dutch. Rizzi's grammatical analysis explains why English-speaking children omit subjects but not objects (in contrast to German and Dutch children) and also why null subjects in early English are restricted to root contexts.

In a similar vein, Hyams (1994) proposed that early English is a residual topic-drop language; that is, null subjects are licensed in [SPEC, CP], as in German and Dutch. However, null objects, which require licensing by a fronted verb in C (the V-2 effect) are excluded because English is not a V-2 language. According to this analysis, then, the difference between early English, in which only null subjects are licensed, and early German and Dutch, in which both null subjects and null objects are possible, is related to an independent difference in the two grammar types, the V-2 phenomenon. In short, a grammatical approach to early cross-linguistic variation assumes that the differences that exist between child grammars (like the differences that exist between adult grammars) fall within well-defined limits. Thus, the goal is to: (a) tease apart the universal aspects of early grammar from those properties of early grammar that are fixed by experience with a particular input language, and (b) determine the limits of the language-particular variation.

In contrast to the grammatical approach, there is the view that cross-linguistic differences such as those that we find with respect to null arguments reflect the operation of different underlying cognitive mechanisms. Such a

---

5 See also Jaeggli and Hyams (1988) and Hyams and Wexler (1993) for earlier related proposals concerning the subject-object asymmetry in English. More recently, Sano and Hyams (1994) argued that the null subject stage in early English is directly tied to the "root infinitive" phenomenon (see Wexler, 1992) and is thus distinct from both Italian pro-drop and Germanic topic-drop.

---

12. COMNEXT: NULL SUBJECTS IN CHILD LANGUAGE

position is illustrated in what I understand to be Valian's proposal: that Italian children omit subjects as a grammatical option (Italian is a null subject language), whereas English-speaking children omit subjects due to performance limitations (cf. also P. Bloom, 1990). Let me refer to this as the performance approach.

In this commentary I argue that the performance approach to the null subject phenomenon is empirically inadequate and theoretically unmotivated. It fails to account for many of the basic statistical properties associated with null subjects in early English and it does not explain the cross-linguistic differences that it sets out to explain. The grammatical model fares better in both these respects. My discussion of the performance account is based on work that I have done in collaboration with Wexler (Hyams & Wexler, 1993).

A second related question that I address is whether the Null Subject Parameter comes preset at an initial default setting, as proposed, for example, in Hyams (1983, 1996). Both Lillo-Martin and Valian find the assumption of an initial [+null subject] setting problematic. Lillo-Martin suggests that the English-speaking child may have different settings for different domains (i.e., root vs. embedded clauses). Valian proposes that the child initially entertains both a [+null subject] and a [−null subject] grammar, "until sufficient evidence accrues to favor one over the other." I compare the standard parameter-setting model to the scale model proposed by Valian.

THE INADEQUACIES OF A PERFORMANCE ACCOUNT

Performance accounts of the null subject phenomenon in child language (P. Bloom, 1990; Valian, 1991, chapter 11 in this volume) make two important assumptions. The first is that null subjects are not a grammatical option for young English-speaking children and hence do not appear in the grammatical representation of the sentence. Rather, sentence subjects are grammatically represented as either full NPs such as John (henceforth I refer to these as lexical subjects) or pronouns. The subject, whether lexical or pronominal, is subsequently dropped during the production of the sentence because of a constraint on output. A second crucial assumption of this approach, made explicit by P. Bloom, is that lexical subjects such as John impose a greater

---

5 Lillo-Martin objects to the initial [+null subject] setting on learnability grounds: the child should not make the assumption that null pronouns are allowed until he or she has positive evidence in the form of sentences with null subjects. However, the learnability problem arises only if the two values of the Null Subject Parameter generate languages that fall into a Subset relation. Hyams (1986) and others have shown that this is not necessarily the case. Depending on the formulation of the Null Subject Parameter, there is positive evidence available that would force the English-speaking child to abandon the initial null subject grammar. I return to this issue.
processing load than pronouns, and that omitting the subject imposes the least load. Thus, the probability of omission is a function of the “heaviness” of the subject selected. So lexical subjects are more likely to be omitted than pronoun subjects.

Hyams and Wexler (1993) developed a formal processing model incorporating these two assumptions, and we tested the model against a number of statistical properties of the null subject stage in early English. The processing model suffers several conceptual and empirical problems. First, it is simply unclear how dropping the subject reduces computational load, because the grammatical representation of the sentence (which, recall, contains a subject) is computed prior to the dropping of the subject, which happens during production. But even allowing that omission does reduce computational load for children, the processing model does not tell us why it is the subject that is omitted as opposed to the object, for example. The subject-object asymmetry is the most salient fact about the phenomenon. Table 12.1 reports the relative frequency of subject and object omission for Adam and Eve (CHILDES, MacWhinney & Snow, 1985). Thus, Adam (age 2;5 to 2;8) omits subjects 55% of the time, whereas objects are omitted in obligatory contexts only 7% of the time; Eve drops subjects at a lower rate, but there is still significantly more omission of subjects than of objects.5 (Similar figures are reported in P. Bloom, 1990; and L. Bloom, Miller, & Hood, 1975; see also Valian, 1991, for comparable data on 21 subjects.)

To explain the subject-object asymmetry within the processing framework, some additional assumptions have to be made. One idea, due to P. Bloom (1990), is that the beginning of the sentence is harder to process than the end of the sentence. This assumption, however, is neither theoretically nor empirically motivated. There is no theory of performance from which such a result follows, and the scant empirical data that exists relevant to children’s productive abilities fails to support the claim (Curtiss & Tallal, 1991; Ferreira & Morrison, 1990).4 On the grammatical account, in contrast, the subject-object asymmetry follows without additional stipulation, because the hypothesis is that the English-speaking child (and all children) start out with a grammar that licenses null subjects.

Another important statistical fact concerns the proportion of lexical subjects relative to pronominal subjects during the null subject stage. The grammatical hypothesis makes a precise prediction in this respect. In a null subject language, null subjects are used where pronouns are used in a non–null subject language (i.e., to refer to contextually specified information). Thus, we would expect that as English-speaking children switch from a [null subject] to a [null subject] grammar, there would be a marked increase in the proportion of pronominal subjects. In other words, the grammatical hypothesis predicts a trade-off between null subjects and pronouns over time, with the proportion of lexical subjects remaining roughly constant. Table 12.2 (adapted from Hyams and Wexler, 1993) reports the proportion of lexical subjects and pronoun subjects for Adam and Eve during the relevant stage of development.

As we can see from the columns headed \(P_i\) and \(P_o\), the prediction of the grammatical model is confirmed. There is a steady increase in the proportion of pronouns over time, whereas the proportion of lexical subjects remains roughly constant from Adam 6 to Adam 30, though there are fluctuations. Similar considerations hold for Eve. Notice, however, that this result is directly at odds with the predictions of the processing model. Recall that according to the processing model, lexical subjects are heavier than pronominal subjects and hence more likely to be omitted. We would therefore expect that in the early periods missing subjects would be more likely to result from the dropping of a lexical subject than from the dropping of a pronoun and thus, that once children grow out of the performance limitation and stop dropping subjects, we would see an increase in the proportion of lexical subjects. This prediction is in no way confirmed.

Notice, moreover, that we cannot explain the trade-off between null and pronominal subjects by saying, as Valian does, that there is some independent factor having to do with the difficulty of pronouns (relative to NPs) that makes them less likely to be used at the earlier ages. If this were the case, we would expect to see the same trends in object position that we see in subject position, that is, we should see an increase in pronominal objects over time. Hyams and Wexler showed that this is not the case. The proportion of pronominal to lexical objects remains roughly constant over time. (See Hyams and Wexler for further discussion.) Thus, the performance account fails to capture the relevant proportions of pronominal, null, and lexical NPs in the early stages, whereas the grammatical account makes precisely the right predictions.

\footnote{Valian finds the same result in her English data. She notes that the percentage of (null + pronominal) subjects remains constant across her different age groups (roughly 89%). Thus, as children develop, they tend to replace null subjects with pronominal ones and the rate of lexical NPs remains roughly constant, as the grammatical hypothesis would predict. Valian attributes this result to the young child's difficulty with pronouns, but see the following discussion in the text concerning pronoun difficulty.}
What about the empirical data that seem to support the processing approach? First, there is the finding that the length of the VP varies as a function of the heaviness of the subject. P. Bloom (1990) did an analysis of the spontaneous speech of Adam, Eve, and Sarah and found that their VPs tended to be shortest with lexical subjects, longer with pronouns, and longest when the subject was omitted. These results are given in Fig. 12.1 (from P. Bloom, 1990).

Hyams and Wexler ran the same analysis on the spontaneous speech of Italian adults, that is, adult speakers of a null subject language, and, strikingly, we found the same result. Our results are given in Fig. 12.2. The similarity between the Italian adults and the English-speaking children shows that the VP length result has nothing to do with production constraints but, rather, is associated with some property of null subject languages, perhaps a pragmatic property that I will not speculate on here.⁶

A second major statistical fact offered in support of the processing model is Valian's finding that Italian children omit subjects at a rate of 70%, whereas English-speaking children omit subjects at a rate of 20% to 50% (depending on the study). According to Valian, this result is explained under the hypothesis that English-speaking children are dropping subjects for performance reasons, whereas Italian children are taking advantage of the null subject option of their grammar. But Valian offers no theoretical reason why a performance constraint should yield fewer null subjects than a grammatical null subject option. In fact, the argument based on frequency differences is a spurious one, because neither the grammar nor the production model makes any prediction at all with respect to frequency. Thus, all Valian's analysis shows is that there is some difference between Italian- and English-speaking children with respect to the use of null subjects. It does not speak to the question of where the difference lies.⁷

To sum up, the processing model does not account for the major statistical properties observed during the null subject stage, whereas the grammatical model makes the right predictions with respect to these same properties.

⁶See Hyams and Wexler (1993) for some suggestion as to the pragmatic nature of the VP length effect.

⁷Although Valian reports that American children in her study drop subjects at a rate of 39% (averaging across studies), she asserts that “American children do not freely omit subjects.” It is unclear what Valian’s criteria are for “free subject omission.” If frequency differences are critical, why is the difference in the rate of subject omission between English- and Italian-speaking children (70% vs. 39%) more compelling for Valian than the difference in the rate of null subject use by English-speaking children versus English-speaking adults (50% vs. 0%)? The computation of the null subject rate for English-speaking children usually does not include those null subject sentences that would be grammatical in the adult language, e.g., nonna le va? Hence, for the purposes of this comparison, adult null subject use is 0%. (Valian, 1991, gives the figure, 0%-10%.)
THE INITIAL STATE: TRIGGERS OR SCALES?

The grammatical approach to children’s null subjects (e.g., Hyams, 1986, in press; Lillo-Martin, this volume, chapter 13; Pierce, 1989; Rizzi, this volume, chapter 10) holds that the initial grammar is a null subject grammar. For children whose target language is not a null subject language, the switch to a [–null subject] grammar is triggered by input data that are not analyzable by the child’s current null subject grammar. For example, in Hyams (1986) I proposed that lexical expletives such as it and existential there, or referential pronouns used in pragmatically neutral contexts, could trigger a parametric shift to a non–null subject grammar, because these elements would not be present in a null subject language. In Valian argues against the idea of an initial setting and the trigger hypothesis, more generally. She proposes, instead, that the child “entertains both options on an equal footing until sufficient evidence accrues to favor one over the other.” I refer to this as the scale model, following Valian’s metaphor.

The argument against an initial setting rests on what Valian calls the parsing constraint. That is, the child’s parser (which is parasitic on his or her grammar) cannot analyze input not generated by the child’s current grammar, and hence it cannot in principle analyze the triggers necessary to induce a parametric shift. Thus, in the case at hand, if the initial setting is [–null subject] the English-speaking child with a null subject grammar would not be able to analyze the lexical expletives or the infelicitous referential pronouns and hence would not be able to move to the correct [–null subject] grammar.

This particular criticism of the parameter model appears misguided, however. The parsing constraint that Valian describes does not block development. On the contrary, within parameter theory it is precisely the assumption of a failed parse under some parameter value that triggers the resetting to the other value. As typically conceived, the parameter-setting model views language development as a failure-driven process (Wexler & Culicover, 1980). Thus, as noted earlier, the child progresses from one developmental stage (i.e., grammar) to the next when he or she encounters input data that are unanalyzable by the current grammar. We can also think of this as a parsing problem; that is, the child’s parser (which incorporates a grammar) is unable to assign a well-formed representation to some input. A failed parse results and the relevant input is said to trigger a change from grammatical state A to grammatical state B. We can formally define trigger as follows (based on Clark & Roberts, 1991):

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

Although the parsing constraint poses no difficulty for the grammatical approach, it raises a vexing problem for the scale hypothesis. Valian argues that the child cannot use as triggering data any input that is not generated by his or her current grammar, that is, any input that results in a failed parse. Thus, the child is caught in an infinite loop, with a parser that is unable to analyze any input beyond the current grammatical stage and a grammar that depends on the analysis of new data in order to develop. The scale model “solves” this problem by providing the child with both grammars/parsers. Armed with both grammars/parsers, the child can parse all of the input data.

---

Footnotes:
8 The precise grammatical characterization of the null subject option—whether pro-drop (Hyams, 1986), topic-drop (Jaeggli & Hyams, 1988), diary drop (Hyams, 1994; Rizzi, this volume, chapter 10), or big PRO, as argued in Sano and Hyams (1994) and elsewhere—is irrelevant to the present discussion.
9 What the particular triggers are depends on the precise grammatical characterization of the null subject phenomenon and is irrelevant to the present discussion, which is intended simply to illustrate the logic of the trigger hypothesis.
10 Kim (1995), in his reply to Valian (1990), made a similar point concerning the parsing constraint. See his paper for a defense of parameter-setting models and arguments against Valian’s hypothesis-testing model, which was very close to the current scale model.
necessary to determine whether the target is a null subject language or not (sentences with and without null subjects, lexical expletives, etc.). But the child at the initial state is faced with not one but many parametric options. Pursuing the logic of Valian's argument, it follows that in order for the child to parse all the triggering data for all parameters, he must start out with the entire set of possible adult grammars (null subject, non-null subject, head-first, head-last, verb raising, non-verb raising, etc.)—a rather implausible assumption, on the face of it.

Not only is the scale model conceptually flawed, but it also engenders serious learnability and computational problems. First, with respect to learnability, if the child starts out with both values of the Null Subject Parameter, he or she has a grammatical analysis for whatever input he or she receives. The child has, in effect, the union of two grammars. On what basis, then, would the child ever reject this "super grammar"? We are faced with a classic subset problem. The Italian child will hear sentences with and without null subjects, all of which, of course, are possible, given his [+null subject] grammar. The English-speaking child, in contrast, will never hear null subject sentences, but he or she could not exclude the null subject option on this basis unless we allow for direct or indirect negative evidence. Thus, adopting the scale model entails that we abandon either the Subset Principle or the assumption of no negative evidence, neither of which is desirable on a priori grounds.

The scale model is also implausible from a computational perspective. On this model, development consists of "gathering," "amassing," "tabulating" evidence in favor of one parameter value or the other until "gradually the evidence weighs down one side very heavily." Clearly, the psychological mechanisms necessary to tip the scales one way or the other require rather large computational resources. At the very least, the child must have an accurate memory for previous input data, past failed hypotheses, and parameter settings, as well as the ability to represent and compare the multiple representations of a sentence that are generated by the competing grammars. Thus, the scale model requires that we abandon yet another well-motivated assumption, namely, that the child is computationally bounded, to use a term of Clark and Roberts (1991). That is, he or she has a finite memory and a (small) finite amount of time in which to converge on the adult grammar. Moreover, the developmental picture that unfolds is one in which the younger you are, the more representations/grammars/parsers you must cope with. See Valian (this volume, chapter 11) for debate on this computational problem.

Quite apart from the plausibility of such a picture, Valian's assumptions concerning the child's computational abilities are clearly at odds with her proposal that the English-speaking child's use of null subjects is a performance effect, as discussed earlier. How is it possible that the child, who has such limited processing resources that he or she is unable to produce sentence subjects reliably, can at the same time compute multiple and in principle very large numbers of representations, for every (structurally ambiguous) sentence?\footnote{According to the assumptions of the scale model, multiple representations would have to be computed only for sentences that are ambiguous (i.e., receive different analyses by the two grammars). However, this restriction does not actually reduce the computational load, because the set of sentences that is structurally ambiguous in this sense is still infinite. Moreover, it is unclear how the grammar/parent would know if a sentence is structurally ambiguous before assigning the structural descriptions.}

In short, the parameter-setting model does not engender the kind of problem Valian discusses. It is a failure-driven model that assures rapid, deterministic learning. The scale model, in contrast, makes acquisition impossible under standard assumptions of a computationally bounded learner without access to negative evidence.

CONCLUSION

An ever-increasing body of cross-linguistic acquisition evidence (of which the papers in this section are a prime example), shows that environmental effects are felt quite early in language development. With respect to the null subject phenomenon, it now appears that Italian children pro-drop, Dutch, German and Chinese speaking children topic-drop (though Germanic topic-drop differs from Chinese topic-drop in certain respects), and English speaking children have yet a different analysis, one possibility being a kind of diary-drop, as Rizzi (this volume, chapter 10) proposes. The finding that children show an early sensitivity to language-specific properties of the input data is a welcome result in that it solves the rather thorny triggering problem of earlier accounts. One of the weaknesses of the parameter-setting analysis in Hyams (1983, 1986), which posited a universal early pro-drop stage, is that it did not explain why children acquiring non-pro-drop grammars fail to analyze the relevant input (triggering) data for as long as they do. Given the current picture, this problem no longer arises in the same way. On the other hand, the range of cross-linguistic variation that we observe very early in development raises the logical problem of language acquisition in a particularly perspicuous way; that is, how do children figure out the particular properties of their language, given such limited and fragmentary experience. Moreover, if we assume, as I think we must, that there is a uniform initial grammatical state (whatever it is) prior to experience, and if it is the case that linguistic input immediately (or almost immediately) alters that state, then the initial state becomes less amenable to direct empirical investigation. In this case, properties of the initial state will have to be determined logically.
rather than empirically, much in the way it is done in the study of adult grammars. We may ultimately find that the instantaneous model of development is closer to the truth than we have imagined.

ACKNOWLEDGMENTS

I would like to thank Robin Clark, Tom Cornell, Kyle Johnson, and Bonnie Schwartz for helpful discussion of a number of issues addressed in this chapter. Thanks also to Ken Wexler for a wonderful collaboration on Hyams and Wexler (1993), on which the first part of this commentary is based. Finally, my appreciation to Jeannette Schaeffer for editorial assistance.

REFERENCES


