

Syntactic subjects in the early speech of American and Italian children*

Virginia Valian

Department of Psychology, Hunter College, 695 Park Avenue, New York, NY 10021, U.S.A.

Received January 10, 1990, final revision accepted January 18, 1991

Abstract

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

Why do young children leave out sentential subjects? Two competence-deficit hypotheses and a performance-limitation account are evaluated in the present set of studies. American children appear to understand that English requires subjects before mean length of utterance (MLU) 2.0. On balance, performance factors account for the data best. Natural conversations between 21 American children (ranging in age from 1;10 to 2;8 and in MLU from 1.53 to 4.38) and their mothers were taped, transcribed, and analyzed to determine when American children understand that English requires subjects. We measured the frequency of subjects (Study 1); types of pronominal subjects, including expletives (Study 2); frequency of modals and semi-auxiliaries (Study 3); frequency of infinitival to, past tense, third person singular, and subordinate clauses (Study 4); length of verb phrase, frequency of different types of verbs, and frequency of direct objects (Study 5). For Studies 1 and 3 we also used, for comparative purposes, transcripts of 5 Italian children, taped monthly for a year. Even our lowest-MLU American group (5 children between 1.5 and 1.99) used subjects and pronominal subjects more than twice as often as the Italian children, and correctly case-marked their subjects. The American children also produced examples of all the sentence elements measured.

*This research was supported in part by a grant from the National Institute of Child Health and Human Development (HD-24369), and in part by a grant from The City University of New York PSC-CUNY Research Award Program. My thanks for helpful comments and criticisms to P. Bloom, R. Fiengo, L. Gerken, J. Katz, W. O'Grady, M. Potter, A. Radford, and the anonymous readers. Rarely can an author have been so fortunate in her reviewers! I am grateful to S. Utakis for her excellent assistance on this project, and to L. Brighenti for her skillful translations and data analysis of the Italian transcripts. My continued thanks to the parents and children whose conversations form the basis of this study.

Introduction

Children's very early speech has long been known for its "telegraphic" character (Brown & Fraser, 1963). Young 2-year-olds leave out, or only inconsistently include, a variety of constituents in their utterances, among them subject noun phrases (NPs), determiners (Dets - "a", "the"), modals ("can", "will", "may", "must"), the copula "be", and so on. Such omissions have received great attention and analysis over the past thirty years, especially in the 1970s (Antinucci & Parisi, 1973; Bloom, 1970; Bloom, Miller, & Hood, 1975; Bowerman, 1973; Braine, 1974; Brown, 1973; Greenfield & Smith, 1976; Valian, 1986 - for a review see Bowerman, 1978). Nevertheless, our understanding of why children's early utterances are short - why constituents are missing - is still rudimentary. The present set of five studies is intended to shed light on this question by focussing on sentence subjects, which are often missing in young children's speech.

There are two basic sorts of explanations of missing constituents: competence explanations and performance explanations. A competence explanation is that children's early grammars do not contain the (adult) rules or structures that would produce utterances containing the necessary elements. A performance explanation is that children operate under processing limitations that restrict the lengths of utterances that they can produce. In the former case, children do not *know* that subject noun phrases, say, are required. In the latter case, children know that they are required, but do not successfully apply their knowledge in every utterance.

Competence and performance explanations need not be mutually exclusive. Both the child's competence and her performance could be deficient. Further, competence and performance factors might operate to different degrees for different constituents. A child could leave out some constituents because she does not know they are required, and leave out others to reduce processing demands. Our principal aim here is to present a fuller picture of children's usage of subjects than has been available from previous work; another is to determine when American children understand that subjects are obligatory and to evaluate competing explanations for children's inconsistent subject use; a third is to examine the interaction of competence and performance factors in acquisition.

It is necessary to assume some linguistic description of the adult grammar that the child is acquiring. Most of the competence explanations that will be examined have assumed a version of modern transformational grammar, called government-binding (GB) theory (Chomsky, 1981, 1982), and that is therefore the formalism adopted here. The formalism is used here both as an approximate description of the adult grammar, and as a constraint on possible hypotheses about the child's early grammar. Thus, descriptions of the child's grammar that would violate assumptions of transformational grammar are not entertained. At the same time, however, we explicitly do not adopt the language acquisition mechanism that is often associated with the formalism (parameter-setting).

In this introduction we first review facts about the nature of subjects in different languages, and then consider different language acquisition accounts. The reader who wishes to have an overview of the findings can read the summary provided at the end of each study.

Language facts and theories

Children's use of subject noun phrases provides a special opportunity to investigate both inconsistent usage of obligatory elements, and acquisition of a fundamental piece of information about one's language – whether or not overt subjects are required. English requires an overt subject in tensed clauses. (The subjects of untensed clauses, or infinitivals, need not be overt, as in "I want to go", where no overt subject appears for "to go". Our discussion will be concerned solely with tensed clauses.)

In English, (1) is grammatical while (2) is not allowed:

- (1) I am a good kid
- (2) * Am a good kid

In contrast, in Italian, both (3) and (4) are allowed (abstracting away from the error of the missing determiner):

- (3) Io sono bravo tato
I am good kid
- (4) Sono bravo tato
Am good kid

There can be an empty or "null" subject in Italian. English and French are among the languages that require overt subjects; Italian and Spanish are among the languages that allow null subjects.

Parameters of language represent linguistically significant dimensions of language, within which there is very narrow variation (Chomsky, 1981, 1982). The null subject parameter refers to the dichotomy between those languages which require overt surface subjects and those which allow null subjects. The null subject parameter is now thought to involve two different parameters: a licensing parameter and an identification parameter (Jaeggli & Safir, 1989; Rizzi, 1986).

The licensing parameter refers to the presence or absence of elements within the inflectional system, called INFL, short for INFLECTION. INFL contains tense and agreement features, and, in English, modals; it is the new-style version of Aux (see Radford, 1988 for background). If the necessary element in INFL is present, null subjects are allowed; if absent, they are not. Exactly what the INFL element is which licenses null subjects is unknown (see, for discussion, Jaeggli & Safir, 1989; Rizzi, 1986). It is the licensing aspect of null subjects which will be our principal concern here.

If a language has the element which licenses null subjects, there are then different possibilities for how the subject will be identified, that is, how the person and number of the subject will be determined. The particular reference of the null subject is thought to reflect some sort of agreement process, though the type of agreement will vary among languages (Jaeggli & Safir, 1989; Rizzi, 1986).

The null subject parameter in principle also controls other language features in addition to whether an overt subject is required (for a list of candidates, see Riemsdijk & Williams, 1986, pp. 298–303). For example, languages that allow null subjects typically do not have “expletive”, or referentially empty, pronouns (though there may be exceptions). An example, in English, of a referentially empty pronoun is the “it” in (5). In Italian, such a non-referential form of “it” does not exist. The only equivalent of (5) in Italian is (6):

- (5) It seems that Jane loves Mary
 (6) Sembra che Jane ama Mary
 Seems that Jane loves Mary

Expletive “it” also occurs in “weather” expressions: “It’s raining”, “It’s snowing”, and so on.

The null subject itself is a form which is abstractly present, but not pronounced. That is, the subject position is not empty: it is filled, but with a form lacking phonetic content. That form has the character of a referential pronoun. Null subjects are usually held to occur only in tensed clauses; the absent subjects of infinitives are considered to be a different empty form – one which is typically anaphoric in character (e.g., Jaeggli & Safir, 1989, but see also Borer, 1989, and Huang, 1989, for different proposals). The null subject of a tensed clause is called *pro* (pronounced “small *pro*” or “little *pro*”), while the null subject of an infinitive is called *PRO* (pronounced “big *PRO*”); see Chomsky (1982) for first mention of *pro*.

Acquisition theories: Two types of competence accounts

Any acquisition theory must specify both a mechanism of grammatical change and the hypothesized nature of the child’s grammar before and after the change. Grammar-based approaches to acquisition are especially useful because they provide details about the child’s formal representations. Such details amount to operational definitions of theoretical claims about a child’s knowledge. For example, the statement that a child has an optional subject cannot be evaluated unless we know what grammatical form that optionality takes; we need to know what mental representation is proposed.

Grammar-based approaches tend to issue in competence-deficit accounts. The empirical analyses we present will bear on two broad classes of competence-deficit accounts. The two differ considerably in their theoretical explanations but overlap to a large extent in their predictions about acquisition.

The pro hypothesis

In one class there is an empty category, *pro*, underlying the superficial absence of subjects (Hyams, 1986, 1987). We will refer to this as the *pro* hypothesis. On the *pro* hypothesis, subjects are not optional: all sentences have subjects, and some of those subjects consist of the unpronounced pronominal *pro*. The child's productions look as if the subject is optional, but there is an underlying abstract subject.

Hyams's initial *pro* hypothesis (1986) provided both a linguistic analysis of the null subject parameter and an account of several features of the acquisition of English and Italian. The core of the account was that all children begin acquisition with the null subject value of the parameter. Children in non-null subject languages like English have to reset the parameter; before they reset the parameter their speech will grammatically resemble the speech of null subject language children, such as Italians.

Hyams (1986) linked the often-noted inconsistent usage of subject noun phrases in American children's early speech to a concurrent absence of expletive subjects, modal verbs, and certain forms of *be*. Consistent subject usage, and use of expletive subjects, modals, and *be*, were predicted to appear at roughly the same time in the child's productions. The mechanism whereby the parameter was reset was the child's coming to notice expletives, which would then force a switch to the other value. Previously collected acquisition data appeared to demonstrate that subjects, expletives, and modals were indeed linked. In a different version of the *pro* hypothesis, Hyams (1987) likens the American child to the Chinese child rather than the Italian child. Since that version is less well developed than the initial *pro* hypothesis, we concentrate on the first version.

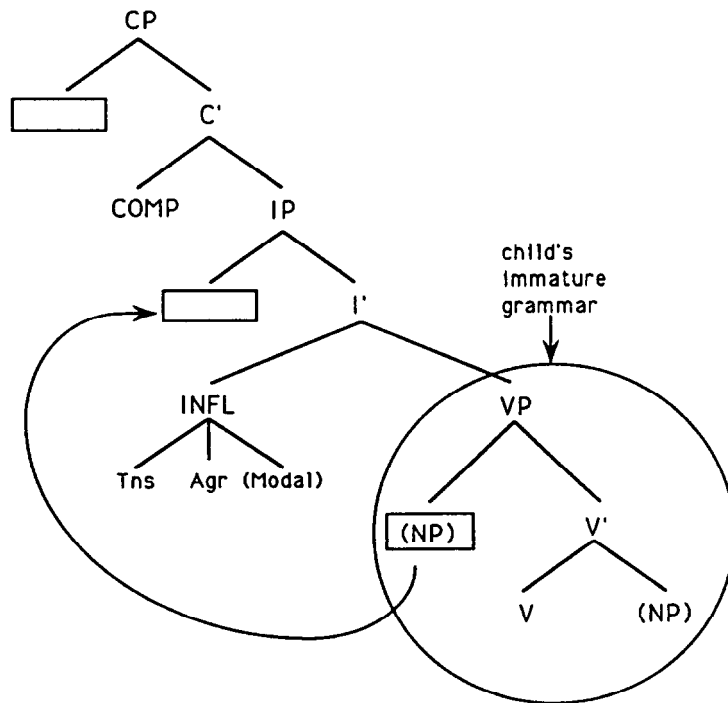
The VP hypothesis

In another class of explanations the child has an immature, or incomplete, grammar (Guilfoyle, 1984; Guilfoyle & Noonan, 1989; Kazman, 1988). We will refer to this as the verb phrase (VP) hypothesis, because the grammar lacks an INFL phrase (and a COMP – short for complementizer – phrase), and consists only of a VP, and the lexical phrases that can appear within a VP (NP, prepositional phrase (PP), and adjective phrase (AP)). Figure 1 shows a tree diagram of the presumed adult representation, with the hypothesized initial child representation circled.¹

For both child and adult, the VP has an optional SPEC (short for specifier) node plus a V' (read as V-bar; see Radford, 1988, for an introduction to X-bar

¹The VP hypothesis has been proposed within the transformational framework. Although it might appear to countenance violations of universal grammar, the spirit of the proposal is that not all parts of universal grammar are immediately available to the child. The grammar only commits sins of omission, not commission.

Figure 1. Tree diagram showing adult English underlying structure and hypothesized child substructure (circled). CP = complement phrase and corresponds to old-style S" (read as S-double bar). Each X phrase has the structure of optional SPEC (specifier) and X' (read as X-bar). SPEC is not itself a category, but a grammatical function, and hence is not shown in the tree. The boxed positions in the tree are SPEC positions. Each X' has the structure of an obligatory head and optional complements. CP thus consists of an optional SPEC and C'; C' consists of the head, COMP, and its complement, IP. IP = INFL (inflection) phrase, and corresponds roughly to old-style S; IP consists of an optional SPEC and I'. I' consists of the head, INFL, and its complement, VP. INFL includes tense and agreement markers, and, optionally, modals; it corresponds roughly to old-style Aux. VP = verb phrase and consists of an optional SPEC, here filled with a "subject" NP, and V'. V' consists of V and optional noun phrase (NP). (The NP notation is used rather than the determiner phrase notation for ease of exposition.) The subject originates under SPEC of VP and moves, in the adult grammar, to SPEC of IP, as shown by the arrow.



syntax) node. VP subjects are inserted under the SPEC node. Since the child's grammar consists of only a VP, and since SPEC is optional, a VP subject may or may not be present. The V' node in turn has an obligatory V node and an optional direct object NP node. On the VP hypothesis there are no genuine sentential subjects in early child speech – only subjects of VPs; those VP subjects are optional. When the remainder of the grammar develops, and INFL is present, genuine subjects which are appropriately cased as nominative can appear.

In the adult grammar for English, the VP subjects will be raised to become subjects of the INFL phrase (as shown by the arrow in Figure 1), and will receive nominative case through a tense or agreement element in INFL. In English, other features of the grammar will force an NP into SPEC of VP. In other languages, raising into IP (INFL phrase) may not be obligatory, because there are other ways in those languages that features in INFL can assign case to subjects. The VP hypothesis predicts absence of tense, infinitival *to*, and modals, and lack of nominative case-marking.

O'Grady, Peters, and Masterson (1989) have proposed that the child initially does not have tense, and therefore cannot distinguish between situations when a subject is required (before tensed verbs) and situations when it is not (before some infinitives). Their linguistic framework, unlike the one we are assuming, allows for sentences without explicit subjects (O'Grady, personal communication). They report that the three children whose corpora they analyzed began using subjects consistently when they began using the past tense productively. Like the VP hypothesis, then, O'Grady et al. predict a relationship between subject use and tense. Unlike the VP hypothesis, however, no relationship is predicted between subject use and modals.

Variable rules

An early account of inconsistent subject use, proposed by Bloom et al. (1975) considerably before the recent work on the null subject parameter, was a model in which the probability of producing a constituent was a function of factors like the familiarity of the verb being used, the inclusion of a negative marker in the utterance, and certain discourse features. Since Bloom et al. say little about the formal representation underlying the children's productions, it is not clear whether the probability model proposes a competence deficit. The formal representation they sketch shows an obligatory subject, thereby seeming to rule out a competence deficit. In addition, recent comments by Bloom (1991) suggest that a performance-oriented model is intended: "children will omit the subject when their cognitive processing abilities are exceeded, for example, when they use new verbs, nouns or pronouns; or add negation or attribution to the sentence".

On the other hand, Bloom, Miller, and Hood (1975) present the model as one which incorporates a variable rule, suggesting a competence deficit. On such a model, the child knows that there are such things as subjects, but does not know

that they are obligatory. The proposal would differ from the pro hypothesis because the abstract element *pro* is not part of the child's grammar. It would differ from the VP hypothesis because the child is not limited to a VP representation. Instead, the child has a full "English" grammar, including genuine subjects, but the child thinks that those subject NPs are optional.

Within the formalism of present-day transformational grammar which we are assuming as a constraint on the child's grammar, a variable rule for subjects is impossible: subjects are mandatory. The only option within our current framework is for the child to believe that subjects are required, but that they can receive a null pronunciation (as may sometimes actually be the case with determiners). A subject NP node is always present, as in Bloom et al.'s (1975) sketch, but the child believes that a subject NP can be morphophonemically empty. The error is not the same as the pro hypothesis, because the abstract element *pro* – with its pronominal properties – is not part of the child's grammar. According to the pro hypothesis the child has an abstract pronominal which lacks phonetic content; according to the no pronunciation interpretation of the variable rules approach, the child inserts no lexical item under the subject NP node. Radford (1990) specifically considers a very similar proposal.

The factors Bloom et al. (1975) propose as determinants for subject use are then factors which determine when the subject NP receives a lexical spell-out: the child's formal representation would or would not lexically spell out the subject NP, depending on the interaction of the performance factors. The factors Bloom et al. investigated, ranging from word familiarity to discourse features, were related in a complicated way to the likelihood of producing an utterance of a given length, as measured by number of major constituents. Most analyses did not specifically concern subjects, but the likelihood of producing two- versus three-constituent sentences.

One important analysis concerned how "complexity" elsewhere in the sentence affected utterance length. Some types of complexity, such as use of a negative, verb + particle construction, or the possessive, were more common in two- than three-constituent utterances. On the other hand, other types, such as the use of verb and noun inflections, modals, semi-Aux's, and some determiners, were as common in three- as two-constituent utterances. Interestingly, verb inflections were more common in subject-verb combinations than in verb-object combinations.

There are no obvious empirical predictions which would distinguish Bloom et al.'s (1975) model, which we may term the "null spell-out" alternative, from a pure performance deficit model. Any findings which support a performance deficit model will also support a null spell-out model. Therefore, even if the other competence-deficit hypotheses can be ruled out, we will not be able to choose between the null spell-out competence deficit and a pure performance deficit.

Others

There are other competence-deficit hypotheses which do not fit so neatly into our two classes, such as Lebeaux (1987), Roeper and Weissenborn (1990), and Radford (1990). Since the major empirical predictions have been made by the pro hypothesis or the VP hypothesis, we concentrate on them here.

*Relation between data and theory: Limitations of competence-deficit accounts**Weak evidential basis*

Competence-deficit theories have been limited by the weakness of the data they have had recourse to. As is proper, acquisition theorists tailor their accounts to fit the reported facts of acquisition. But, surprisingly, given the abundance of theoretical accounts of children's early use of subjects, we lack normative quantitative data on the development of American children's use of subjects, and of related constituents. Existing accounts use either summary data reported in other contexts, or data from very small samples of children (and the Brown corpora have been used repeatedly). It is often difficult to know how common a reported pattern is.²

Before theory development can progress, and before current theories can be winnowed, we need data from a large number of children in a variety of languages (and replications by other investigators with new samples of children) on the quantitative development of subjects and other syntactic elements. We need to know exactly what the phenomenon is that we are trying to account for.

As a first step in that process, we measure American children's use of subjects across a broad mean length of utterance range (MLUs 1.53–4.38), plus their use of the sentential elements which the two principal competence-deficit hypotheses have implicated as linked to knowledge of subjects. We also compare American and Italian children's production of subjects.

²Many authors have relied on data presented in Bellugi (1967), even though those data represent summaries from fairly widely separated time periods for three children (and for some measures data from only one child are presented). A second source is data from Bloom (1970), and Bloom, Lightbown, and Hood (1975). Bloom (1970) includes examples of many types of utterances of 3 children studied longitudinally, and Bloom, Lightbown, and Hood (1975) include in their appendix examples of utterances from Bloom's 3 children plus another child. Presentation of the full corpus for each child (a total of over 27,000 utterances) would obviously have been impossible, and unnecessary for those authors' purposes. But the presentation of examples rather than a systematic sample makes the published data inappropriate as a source of data for analyzing subject use. Finally, Radford (1990) collected 39 cross-sectional samples, and 12 longitudinal samples – a huge corpus. Unfortunately, he presents almost no quantitative data, so his data cannot be compared with ours.

The present study points up the need for investigators to sample large numbers of children, to sample children who are not part of existing databases, and to provide numerical and proportional tabulations of usage of different constituents. Otherwise, the idiosyncrasies of individual children's patterns may be enshrined as typical of language development, and infrequent examples may be accepted as the norm.

The basic fact from which all the theories have started is children's early inconsistent use of subjects. Competence theorists have interpreted this inconsistent usage as a competence deficit: American children do not *know* that English requires explicit subjects, either because they improperly have pro in their grammar (pro hypothesis) or because they have an immature grammar (VP hypothesis). O'Grady et al. (1989) used a criterion of 90% use of subjects to classify children as having an obligatory surface subject. Hitherto, no criterion had been proposed, so that their use of one represented an advance of the field. On the basis of the 90% criterion they concluded that there was a stage of optional subjecthood, since the three American children they examined (two from the Brown corpora) did not start out using subjects 90% of the time.

The problem with the 90% criterion, initially developed by Brown (1973) for other syntactic elements, is that it is probably too strict.³ While it is safe to say that a child who uses subjects in 90% of the required contexts understands that overt subjects are obligatory, it is not equally safe to say that a child who uses subjects less than 90% of the time lacks that knowledge. The lower the child's MLU, and the lower her level of development generally, the more difficult it will be for her to express her knowledge fully, and the more likely it is that her usage will be inconsistent. Inconsistency is not in and of itself a hallmark of a competence deficit. Therefore, we take it as an open question, and a theoretical question, whether the American child ever thinks that overt subjects are not required.

Ancillary assumptions

A second limitation of competence-deficit accounts is that, on some matters, as Hyams (1987) notes, linguistic theory is mute, and ancillary assumptions must be brought in to derive predictions. The example Hyams (1987) gives is of the initial setting of the parameter: linguistic theory is neutral with respect to which value should be the initial value (if any). In certain cases (as with modals, discussed below) predictions are derived from the additional assumptions rather than from the linguistic analysis.

³Brown (1973) reported on the points at which 3 children produced various morphemes in 90% of the contexts in which they were obligatory in 3 successive taping sessions. The criterion has often been used as if it represented knowledge onset. But the children clearly had knowledge of the morphemes in question considerably before they could supply them in 90% of required contexts. For 2 of the children, 5 of the 14 morphemes examined had not reached criterion by MLU 4, and for the third child 8 out of 14 morphemes had not reached criterion by MLU 4. MLU 4 is quite an advanced point in acquisition. We therefore think Brown's morpheme data are an inappropriate source for determining knowledge onset. In addition, given the much larger sample examined by de Villiers and de Villiers (1973), their morpheme data are preferable.

Acquisition theories: Performance accounts

Under the heading of “performance factors” we will include a large variety of variables, representing in some cases limitations (as in limited working memory), and in others merely performance characteristics (as in prosody and discourse factors).

Performance limitations

Children, being human, are variable and work under performance constraints. One production constraint is memory. Young children have a smaller working memory than do adults, and their memory span is correlated with their MLU (Blake, Quartaro, Austin, & Vingilis, 1989). Chi (1978) has claimed that children’s inferior memory is due to their lesser ability to form chunks (rather than to capacity limitations): children can chunk fewer items together into a unit than an adult can, because they have less experience with most domains than adults do. Thus, Chi links lack of knowledge and experience with smaller memory. Olson (1973) attributes children’s poorer memory to “their failure to organize, plan, monitor, and integrate their information processing and remembering as effectively as older children or as adults ...” (p. 151). It is clear that a 2-year-old has much less expertise with the various sequential tasks of talking and listening than an adult has, as well as less grammatical knowledge, and is therefore likely to form smaller chunks than an adult. Children’s lower efficiency will result in their producing shorter utterances overall and shorter constituents within those utterances.

In the case of adult speech we *know* that the length of utterances is due to extra-syntactic performance factors. Our grammars allow us to produce very long utterances. But the combination of planning the content of what we have to say, finding and organizing the syntactic structures to express the content, finding the words, taking into account the listener’s memory limitations, being a good conversational partner, and so on, all conspire to limit the length of our productions. Children have the same tasks, and much less practice at integrating them, so they will have at least as many constraints as adults do. Thus, it is highly likely that children’s performance system is also a limitation on the length of their utterances, and the younger the child the greater the limitation.

It is clear that length limitations alone cannot predict which constituents the child will fail to include. It will be necessary to couple length limitations with other determinants, such as the content of the message the child wants to convey, syntactic requirements, and discourse requirements, to arrive at the systematic omission of function words compared to content words.

Bloom (1970) argued that some of children’s omissions represented reductions of elements that were present in deep structure, not all of which could be expressed in surface structure, because of performance limitations. Although Bloom presented reduction as a transformational rule (which would make it part of

competence), it seems more in keeping with her argumentation to regard reduction as a performance process. Bloom (1990) has presented data from three children (the Brown corpus) supporting a performance explanation for children's early inconsistent use of subjects. The children's verb phrases were longer when a subject was absent than when it was present, as would be expected if children were operating under performance limitations.

There are also data suggesting that the child's processing load is higher at the beginning of an utterance than at the end. NPs consisting of a determiner, adjective, and noun are infrequent in children between MLUs 2.93 and 4.14, and occur only as objects for children below MLU 3.5 (Valian, 1986). That asymmetry suggests that objects are easier to elaborate than subjects, perhaps because planning the utterance is more effortful at the beginning of the utterance (Pinker, 1984), and perhaps because English is right-branching (Lust & Chien, 1984). Mazuka, Lust, Wakayama, and Snyder (1986) also invoke processing factors. On the basis of Japanese children's data, as well as reports of English-speaking children's data, they analyze children's omission of subjects as due to sensitivity to the principal branching direction of their language plus processing constraints.

It is, however, also likely that pragmatic factors play a role in the distribution of NP types. Any NP containing a noun is more common as an object than as a subject at every MLU observed, while NPs consisting of a pronoun are more common as a subject than as an object (Bloom, 1990; Limber, 1976; Valian, 1986. Bloom, Lightbown & Hood, 1975, found this pattern to be somewhat variable at low MLUs). The heavy use of pronominal subject NPs may thus reflect two factors: first, that the subject tends to be "given", or old, information, while the object tends to be new information; second, that processing demands are higher at the beginning of an utterance than near the end, and pronouns are the least taxing NPs one can produce.

Unfortunately, it is difficult to derive explicit predictions from performance accounts. L. Bloom's and P. Bloom's predictions are one set. We attempt in Study 5 (on verbs and direct objects) to develop and test additional performance predictions.

Acceptability

Adult English speakers omit the subject in a variety of contexts (O'Grady et al., 1989; Valian, 1990a, 1990b). Expletive "it" is omitted, as in (7), from the *New York Times*, as are other subjects in extended discourse:

- (7) Seems like she always has something twin-related perking

The discourse and prosodic conditions on subject omission in English usage are quite complex (Valian, 1990a). Omission is more acceptable before a main verb than before a modal, as is shown by the greater acceptability of (9) compared to (10) as a continuation of (8):

- (8) She'll be a big hit
- (9) Sings like a dream
- (10) Can sing like a dream

Similarly, omission is more acceptable before a main verb than before auxiliary *be*, as is illustrated by the contrast between (12) and (13) as continuations of (11):

- (11) She's going to be a big hit
- (12) Sings like a dream
- (13) Is singing like a dream

Yet when the modal or *be* is concatenated with a negative, the string becomes more acceptable, as shown by the sequence in (14) and (15):

- (14) She'll never make the big time
- (15) Can't sing worth a nickel

What is acceptable in the adult community forms part of the child's input, and is also part of what children must master. The utterances that I have termed "acceptable" are not grammatical in English (since English does not have *pro* subjects, and also cannot be characterized as a simple VP). They lack subjects and therefore violate the extended projection principle (Chomsky, 1981), which we are assuming.

Children are exposed to fully grammatical utterances without subjects, in the form of imperatives. They are also exposed to acceptable utterances which are not fully grammatical, such as (9), as well as forms like, "Want lunch now?" The American child must grow into an adult who not only knows that overt subjects are grammatically required, but also knows when subjects can acceptably be omitted. The child must not only acquire the correct grammar, but also master the discourse conditions that allow relaxation of the grammar.

Utterances without subjects could be difficult input for the child in two ways. The child could be misled into thinking that subjects are not grammatically required (O'Grady et al., 1989; Valian, 1990a, 1990b). Even if the child has determined that subjects are grammatically required she may not have learned where subjects can acceptably be omitted. Given how complicated the conditions on acceptable omission are, it will take the child time to zero in on the correct contexts – how much time is not clear. The result in either case in the child's speech would be subjects which were absent too frequently: in the former situation the child's competence would be deficient, in the second the child's performance. If children know subjects are obligatory, they should confine omissions to linguistic contexts where adults might also omit them. Thus, in Study 3 we contrast children's use of subjects before modals versus main verbs, and in Study 5 we measure subjects in subordinate clauses (where subjects cannot be acceptably omitted).

Another way of conceptualizing “acceptable” utterances is to think of them as grammatical within a different grammar. From that perspective, adults have a variety of grammars, and use different ones in different discourse conditions. But there are problems with that conceptualization. One is the need to posit that adult English speakers have at least two grammars: one the grammar that we usually think of as holding for English, another that either contains *pro* or allows for VP clauses. The second is the possibility that an indefinite number of grammars would then be attributed to any adult, in order to account for all the non-standard utterances that are produced. In any event, on this conceptualization as well, the child’s task is more complicated, since she must acquire more than one grammar, and sort out the principles underlying each.

Prosody

Gerken (1991) suggests a prosodic explanation of the child’s tendency to omit unstressed syllables in utterance-initial position, which can also be applied to adult speech phenomena. If a bias exists against beginning an utterance with an unstressed syllable, that unstressed syllable will either be assimilated onto a following stressed syllable, or deleted altogether. The lesser acceptability of (10) compared to (9) may thus have nothing to do with the distinction between modals and main verbs, but may only reflect our preference to begin sentences with a stressed syllable; (10) does not lend itself to initial stress. The “weak” nature of modals and *be* may make them less acceptable in initial position; when they are concatenated with a negative, they become “stronger”, and thus more acceptable.

Gerken (1991) had young 2-year-olds imitate sentences with different types of subject and object NPs (consisting of proper names, pronouns, or full lexical NPs). She found that children omitted subject NPs more often than object NPs – roughly 19% omissions versus less than 1% omissions. Children also omitted determiners more often in subject than object NPs, roughly 31% omissions versus 18% omissions. Pronoun subjects were omitted about 32% of the time, proper nouns and common nouns about 12% of the time. Gerken hypothesizes that initial weak syllables of a metrical foot have a certain tendency to be omitted, so that the same prosodic process is responsible for the omission of subject pronouns and the omission of subject determiners. In addition, Gerken reports that when a determiner in object position is the initial syllable of a metrical foot, the omission rate is the same as when it is in subject position. That would suggest that omission of initial weak syllables, regardless of sentence position, is responsible for subject omission.

There are, then, ample reasons for thinking that factors independent of the child’s syntactic competence are partially or wholly responsible for children’s omission of subjects in particular and of syntactic elements in general. Those factors may include processing load, characteristics of the input, ignorance of the exact acceptability conditions for subject omission, and prosodic effects. In all these cases, subjects are more prone to reduction or omission than objects.

Teasing apart competence and performance

The present study uses two strategies to tease apart competence and performance factors. One is to explore both competence and performance factors within English. To begin with the American children, we examine properties that the children's speech should have, on different theories, if they understand that subjects are obligatory, using as a benchmark the speech of children who are consistently using subjects. Thus, we measure use of: subjects, pronominal subjects, nominative case-marking of pronominal subjects: expletives; modals: infinitival *to*, past tense, and third person singular; subordinate clauses; direct objects. Simultaneously, we look for evidence of performance limitations.

The use of a benchmark, in the form of the speech of children who are consistently using subjects, is an important control. If children who use subjects consistently use, say, few expletives, the absence of expletives in the speech of children who use subjects inconsistently is not informative.

The second strategy is to compare English-speaking children with children learning a null subject language, such as Italian. If Italian children are used as a benchmark, American children's similarity to them can be assessed. We compare production of some of the same elements in the two languages, such as subjects, pronominal subjects, and modals. If the measurements yield the same results, we can conclude (a) that American children are not sensitive to the differences between their language and a null subject language or (b) that the measures reflect universal performance characteristics. If the measurements yield different results, we can conclude that American children are sensitive to how subjects are used in their linguistic community.

Study 1: Subject use

The first and most basic question is how frequently children use subjects. Frequency figures alone will not tell us whether children understand that their language does, or does not, require subjects, but the facts about usage will tell us what the phenomena are that need explaining. Data from American and Italian children will allow us to contrast usage between speakers of a non-null subject language and a null subject language, to see whether differences between those two canonical language types are reflected in the speech of children, or whether children treat the two languages as if they were the same.

Competence-deficit predictions

On the pro hypothesis (Hyams, 1986, 1987), the American child, like the Italian child, has available three possible types of subjects: full lexical noun phrases (NPs), pronoun NPs, and pro NPs. If we ignore temporarily the possible effects

of different input from adults of the two languages, the children of both languages should produce null subjects, and produce them equally often, since they both think their language is a null subject language.

Given that a pro NP is considered to be an alternative version of a pronoun NP, we can also derive the prediction that there will be fewer pronoun NPs in the speech of a null subject language speaker than in a non-null subject language speaker. This is especially so if, as Hyams (1986) hypothesizes, there is a principle like avoid pronoun (Chomsky, 1981), which states that the speaker will not use a pronoun NP, but a pro NP, as the default in a null subject language. Thus, if children believe their language is a null subject language, and if their performance is constrained by the avoid pronoun principle, there should be few pronominal subjects in their speech. The subjects that are expressed should be primarily lexical rather than pronominal. That prediction can be verified by looking at pronominal NPs in American and Italian children's speech.

On the VP hypothesis (e.g., Guilfoyle & Noonan, 1989), children of all languages should look the same at an early stage of development: their grammars consist of VPs, with optional subjects. If American children at an early point in development (say, under MLU 2.0) have only VP grammars, then they should be indistinguishable from children of other languages at a comparable developmental point, again ignoring possible effects of input.

Neither the pro nor the VP hypothesis specifies any particular role for input. Both make clear claims about the form of the child's grammar, but not about the child's output. The child's grammar is one determinant of her output; her performance system is another. If those were the only two determinants, then we could confidently predict equal use of subjects by all children regardless of target language. But, since there is evidence that children's frequency profiles match their parents' (Brown, Cazden, & Bellugi, 1973), we might expect parental use of subjects to influence child usage. American and Italian adults do use subjects to different degrees, and we might thus expect differences in children's speech on that basis alone. We will return to this problem in the discussion.

Performance predictions

Like the competence-deficit accounts, performance accounts also make no firm predictions about amount of subject usage, but a performance account will predict certain differences between American and Italian speech. If the American child understands that her language requires subjects, and omits subjects for performance considerations only, there should be more subjects in her speech than in the speech of an Italian child who understands that overt subjects are syntactically optional in her language. Similarly, there will be no syntactic reason for American children to omit pronouns if pro is not part of their grammar. Therefore, a performance account should also predict a higher degree of pronoun usage among Americans than Italians.

Method

American corpora

Children and taping procedures. Twenty-one children (12 girls, 9 boys) were audiotaped, using a cassette tape recorder, in natural conversation and play with their mothers. Two-year-olds were recruited among acquaintances of the author and her assistants, and via bulletin boards and local newspaper advertisements. The children ranged in age from 1;10 to 2;8. All the children were white, and the socio-economic status of their parents ranged from working class to upper middle class. Six households had one or two Ph.D.s or graduate students.

Typically there were two taping sessions per child–parent pair, no more than 2 weeks apart. Session 1 was usually half an hour, and Session 2 one hour. Taping was performed at the child's home or day care center, or in a college play room.

Transcription and MLU calculation. Each tape was transcribed by one listener and checked by another, with the exception of a child of MLU 3.72, whose transcript came from a previous study. Conventional English orthography was used. Doubtful portions were placed in parentheses, while completely unintelligible portions were so indicated.

The children's MLUs in morphemes were calculated according to Brown's (1973) procedures for each taping session. Each child's average MLU was then used as the most reliable estimate of the child's overall linguistic level during the taping period. The MLUs ranged, roughly evenly, from 1.53 to 4.38. For this study, the children were divided into four groups, based on their MLUs. Table 1 shows the distribution of children by age and MLU.

Observers were asked to note clarifying context, in particular so that imperatives could be distinguished from declaratives which were lacking a subject. This was done for the purposes of Valian (in press), where the grammaticality of children's utterances was coded. (An imperative would be classified as grammatical, but a declarative missing a subject would be classified as ungrammatical.) For the higher MLU children (above 3.0) the surrounding verbal context was usually sufficient to determine whether an imperative was being spoken, but for the lower MLU children (under 3.0) the observers' clarifications were helpful. The present study was not contemplated until most of the grammaticality assignments had been made. It is therefore unlikely that there was any theoretical bias in deciding whether an utterance was an imperative.

Utterances used in analyses. Three classes of child utterances were eliminated from analysis. The first class, discards, consisted of utterances unintelligible in whole or in part, interrupted utterances, and utterances that consisted solely of single-word assents, dissents, or hesitations. (Unintelligible utterances, interrupted utterances, and hesitations were also excluded from MLU calculations.)

Table 1. *American children: MLU, age, and utterance information*

	MLU	Age	Number total utterances ^a	Number + V utterances ^b	Proportion + V utterances
Group I	1.53	2;1	474	47	.10
	1.74	1;10	491	169	.34
	1.79	2;2	322	76	.24
	1.81	1;10	259	62	.24
	1.99	2;2	187	76	.41
Mean	1.77	2;0	347	86	.27
Group II	2.24	2;3	264	106	.40
	2.28	2;3	442	227	.51
	2.52	2;7	566	274	.48
	2.66	2;8	330	182	.55
	2.76	2;5	390	250	.64
Mean	2.49	2;5	398	207	.52
Group III	3.07	2;3	534	406	.76
	3.15	2;6	285	217	.76
	3.16	2;3	520	287	.55
	3.31	2;5	418	289	.69
	3.44	2;5	437	322	.74
	3.62	2;5	441	309	.70
	3.68	2;5	359	275	.77
	3.72	2;5	291	182	.63
Mean	3.39	2;5	411	286	.70
Group IV	4.12	2;8	268	225	.84
	4.17	2;8	333	251	.75
	4.38	2;6	347	273	.79
Mean	4.22	2;7	316	250	.79

^aNumber of non-imperative non-imitative utterances; this is the denominator for calculating the proportion of total utterances with a subject.

^bNumber of non-imperative non-imitative utterances containing a verb; this is the denominator for calculating the proportion of +V utterances with a subject.

The discarded class was roughly 40–50% of the child's utterance tokens. Utterances with doubtful portions were included in usable utterances. Repetitions of whole utterances were counted as separate utterances; we counted utterance tokens, not types.

Since the focus of the study was the child's spontaneous productions, we subtracted two other classes from the set of usable utterances: imitations and stock, routine utterances. An imitation was an utterance which directly followed an adult utterance, was either a full or partial repetition of that utterance, added no new material, and made no morphemic alterations (Ervin, 1964). There was one

exception to this. If the child repeated a parental utterance that was itself an imitation of the child's spontaneous preceding utterance, the child's repetition was not scored as an imitation. A routine was a stock utterance type used repeatedly by the child with no variation from one time to the next (cf. Brown, 1973; Peters, 1983). Table 1 shows the number of usable utterances per child, which ranged from 187 to 566.

Questions as well as declaratives were included in the analyses (in both cases only if the utterance included a verb). *Wh*-questions required a scoring decision. In the adult grammar, most of the *wh*-words would not be considered subjects, but (underlying) objects or adjuncts of the verb (moved into COMP). The question "What is that?" would be considered to have "that" as its subject in the adult grammar, not "what". Of the *wh*-words the children used, *what* and *who* could be (underlying) subjects or objects in the adult grammar, depending on the sentence; "What happened?", for example, would be considered to have "what" as its underlying subject in the adult grammar. *Where*, *how*, and *why* would be classified as objects or adjuncts.

In the child grammar, we do not know how the child is analyzing questions like "What is that?" For the child "what" could incorrectly be base-generated as the subject, or as the object moved into COMP, or as the object moved into some other position. We included such question tokens as part of the children's data for two reasons. First, in Italian, it is possible grammatically to say the equivalent of "what is?", or "where is?" If American children include both a *wh*-word and a "subject" NP, that is noteworthy. In fact, almost never did a child omit the true subject: out of a total of 552 *wh*-questions where the *wh*-word would not be classified as the subject in the adult grammar, the children supplied a subject in all but 9 cases, evenly scattered among the MLUs. Second, for all the children (except MLU 1.81, who produced no *wh*-questions), their questions appeared at least minimally productive. Even the lowest-MLU child, whose only *wh*-word was "what", and whose only accompanying verb was the copula "is", showed some productivity, using "it", "this", and "that" as the "subject", using both the contracted and uncontracted form of "is", and once using no verb.

For the subject counts, then, it did not matter whether in the child's grammar the *wh*-word or the other NP was classified as the subject, because the children included both. We did, however, also perform an analysis excluding *wh*-questions, to see whether that would reduce the children's percentage of subjects. In counting the number of pronominal subjects, we performed two analyses: one in which we included *wh*-questions and one in which we excluded them. When we included *wh*-questions we decided to err on the side of conservatism, and scored *what* and *who* as the subject. In almost all such cases, the true subject was itself a pronoun (usually "this", "that", or "it"), and thus a pronominal subject would have been scored either way.

Italian corpora

Children, taping, and transcription. Five northern Italian children were audio-taped in conversation with an observer. Transcription was performed by G. Tirondola, and the transcripts were lent to the author by F. Antinucci. The children were observed 11 times, beginning at age 1;6 or 1;7. The observations occurred once a month, except for a two-month hiatus encompassing a summer vacation; the break occurred during Month 6.

From the length of the transcripts, each recording session appears to have lasted no longer than 15–30 minutes. For that reason, it was necessary to pool sessions. We divided the data into Time I and Time II. Time I covers the first 5 sessions, before the summer break; the children were about 1;6 to 1;10 during that time. Time II covers the last 6 sessions, after the summer break; the children were 2;0 to 2;5 during that time. Table 2 shows the number of usable utterances per child, and the number of utterances containing a verb per child. Since MLU cannot be calculated in the same way with Italian children as with American children, it was not computed.⁴

A native Milanese, who was also a graduate student in linguistics in New York, and who speaks English fluently, translated the transcripts into English, and coded them. The author reviewed the translations and coding in discussions with the translator.

Utterances used in analyses. In discarding utterances we used the same criteria that we used with the American children. Almost none of the utterances were unintelligible; we surmise that unintelligible utterances were not transcribed. From the set of usable utterances we subtracted imitations and routines. We did not separate declaratives and questions since there were so few utterances available per child. For the 5 months comprising Time I there was an average of 150 usable utterances per child; for the 6 months comprising Time II there was an average of 202 usable utterances per child.

Comparison of American and Italian children

Age. The American children in Group I ranged from 1;10 to 2;2, with an average age of 2;0. The Italian children at Time I ranged from 1;6 to 1;10, and

⁴English has relatively little morphological complexity compared with Italian. In American children's speech bound morphemes are largely confined to plural *-s*, third person singular present tense *-s*, progressive *-ing*, and past tense *-ed*. Italian children have gender and number for nouns and adjectives, person, number, and tense for verbs, and some preposition–determiner combinations. It is difficult to know when these morphemes are genuinely analyzed as morphemes by the children. If they are assumed to be analyzed from the beginning, the Italian child's MLU is very high relative to the American child's.

Table 2. *Italian children: utterance information*

	Child	Number total utterances ^a	Number + V utterances ^b	Proportion + V utterances
Time I	E	119	43	.36
	O	133	32	.24
	M	168	42	.25
	G	139	41	.29
	D	192	39	.20
Mean		150	39	.27
Time II	E	164	55	.34
	O	196	83	.42
	M	219	86	.39
	G	211	84	.40
	D	220	84	.38
Mean		202	78	.39

Note. Children were aged 1;6 or 1;7 at start of taping. Children were taped once monthly, except for Month 6. Time I comprises Months 1-5; Time II comprises Months 7-12.

^aThe number of non-imperative non-imitative utterances: this is the denominator for calculating the proportion of total utterances with a subject.

^bThe number of non-imperative non-imitative utterances containing a verb: the denominator for calculating the proportion of +V utterances with a subject.

at Time II from 2;0 to 2;5. The Italian children were thus younger than the American Group I at Time I and older at Time II. They were slightly younger than Group II at Time II.

Verb usage. In order to have a linguistic way of comparing the children, we calculated the proportion of non-imitative, non-imperative utterances with verbs for all groups of children. The proportions for American children are shown in Table 1, and those for Italian children in Table 2. The Italian children at Time I produced the same proportion of utterances with verbs as the American children in Group I (.27), and at Time II produced a higher proportion (.39).⁵

⁵The present study compares cross-sectional data from Americans with longitudinal data from Italians. It is necessary to ask whether cross-sectional data can be used to investigate children's use of subjects. It could have turned out that there was so much variation among children in their rate of production of subject noun phrases and other constituents that a cross-sectional study based on either age or MLU would be unrevealing. If subject use and related phenomena are unrelated to age or MLU, then cross-sections made on the basis of age or MLU will not show an orderly development. As will be seen, however, the data reported here are orderly. On most measures there is a relation with MLU, the primary basis on which the children were grouped. Nevertheless, future work should also be undertaken to confirm our cross-sectional data with longitudinal analyses. One advantage of cross-sectional data should also be noted: a larger sample of children is thereby possible.

Measures of subject use

We measured subject use two ways: (1) total appearance in all usable non-imitative, non-imperative utterances, and (2) appearance within utterances containing a verb.

Measure 1: Total subject appearance. The denominator consisted of all non-imitative non-imperative usable utterances, including questions (shown in Table 1). The numerator consisted of utterances containing verbs which also contained a subject noun phrase (as in Measure 2 below), as well as utterances consisting of a subject NP plus an object NP, with a missing verb, typically a copula. Examples of strings classified as having a subject despite verb absence are: "this a good story", "birdie inside", "eggs in house", "this outside", "hand in there?", "it noise", "yeah, I my diaper on now". The parent's response was used to resolve difficult cases: "Baba [referring to the child] popcorn, too" was scored as containing a subject; the parent's response was "you like popcorn, too". *Wh*-questions lacking a verb but containing a *wh*-word plus object were included in the numerator.

Cases where neither the form nor the parental response justified coding a subject were not included. A fragment consisting of a single noun phrase, for example, is impossible to categorize as subject or object NP, and would be included only in the denominator. Cases of a noun followed by an adjective were typically scored as NPs rather than as a noun and a predicate adjective. As a result, the complement of subject appearance cannot be interpreted as subject omission.

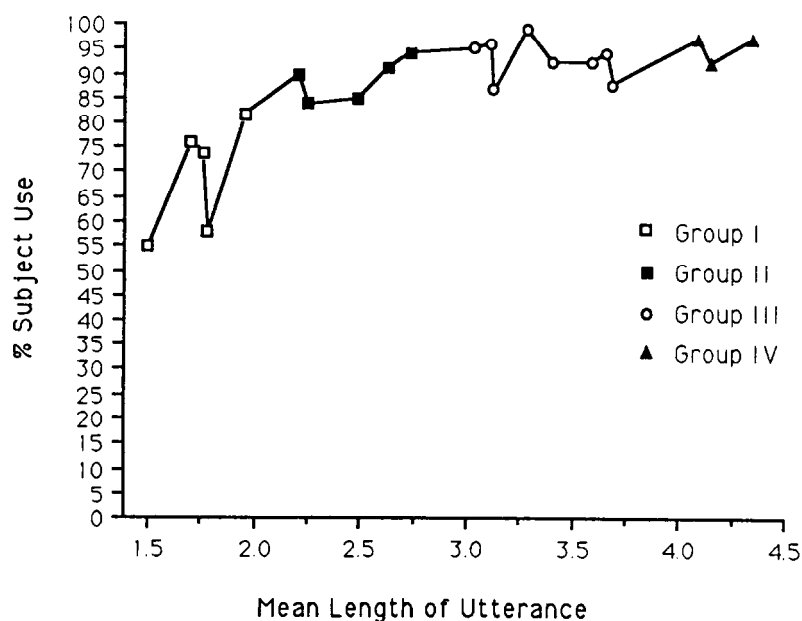
Measure 2: Subject appearance in utterances with verbs. The denominator consisted of all non-imitative, non-imperative usable utterances with verbs, including questions (shown in Table 1). The numerator consisted of the subset containing subjects. We also performed an analysis on the American children excluding *wh*-questions. In Measure 2, the complement of subject appearance *can* be interpreted as subject omission.

Measures of pronominal subject use

We had two measures of children's use of subject pronouns: (1) total subject pronoun use, and (2) for American children only, total subject pronouns excluding *wh*-questions.

Measure 1: Total subject pronoun use. The denominator consisted of all utterances including a verb and including a subject (the numerator from the measure in Figure 2). The numerator consisted of all subject pronouns in those utterances, including personal pronouns, demonstrative pronouns ("this", "that", etc.), and the interrogative pronouns "who" and "what". Variants of "What's this?" (such

Figure 2. American children's use of subjects in non-imitative non-imperative utterances with verbs.



as "What is this?", "What's that?") were scored as having an interrogative pronominal subject. Only utterances with verbs were included.

Measure 2: Subject pronoun use, excluding interrogatives. To ensure that *wh*-questions were not skewing the results, they were excluded from Measure 2. The denominator thus consisted of all utterances including a verb and including a subject, minus *wh*-questions. The numerator consisted of all subject pronouns in those utterances, including personal and demonstrative pronouns.

Results

We first present the American and then the Italian data.

American children

Subject use. Measure 1 of subject use was the percentage of subjects in all non-imitative non-imperative usable utterances, with or without a verb. In Group I the mean percentage is very low – 23%; it increases to 50% in Group II, 68% in Group III and 77% in Group IV.

Figure 2 plots the results for Measure 2 of subject use (percentage of subjects in non-imitative non-imperative usable utterances with verbs), for each child individually. The MLU group into which each child was placed is shown in the legend.

Figure 3 shows the means of subject use (Measure 2) for each MLU group. If one concentrates on the bars for American children, one sees that the 5 children in Group I showed an average use of subjects of 69% (s.d. = 12). There is a clear increase in per cent usage between Group I and Group II: the 5 children in Group II showed an average use of 89% (s.d. = 4). Performance was consistently high thereafter, increasing slightly in the next two groups (Group III mean = 93% (s.d. = 4); Group IV mean = 95% (s.d. = 3)). As the standard deviations indicate, Group I was the most variable group. Table 3 presents the data individually for the children in Groups I and II. An important question is the significance of the 69% figure for Group I, in light of the variability in this group, where two children show subject use as low as 55% and 58%.

When *wh*-questions were excluded, the results were almost identical except for two children in Group I. For the lowest-MLU child, subject use declined from 55% to 38%, and for the child at MLU 1.74 the decline was from 76% to 69%. That reduced the average subject use for Group I from 69% to 64% (s.d. = 17).

Figure 3. Comparison of American and Italian children's production of subjects and pronominal subjects in utterances with verbs.

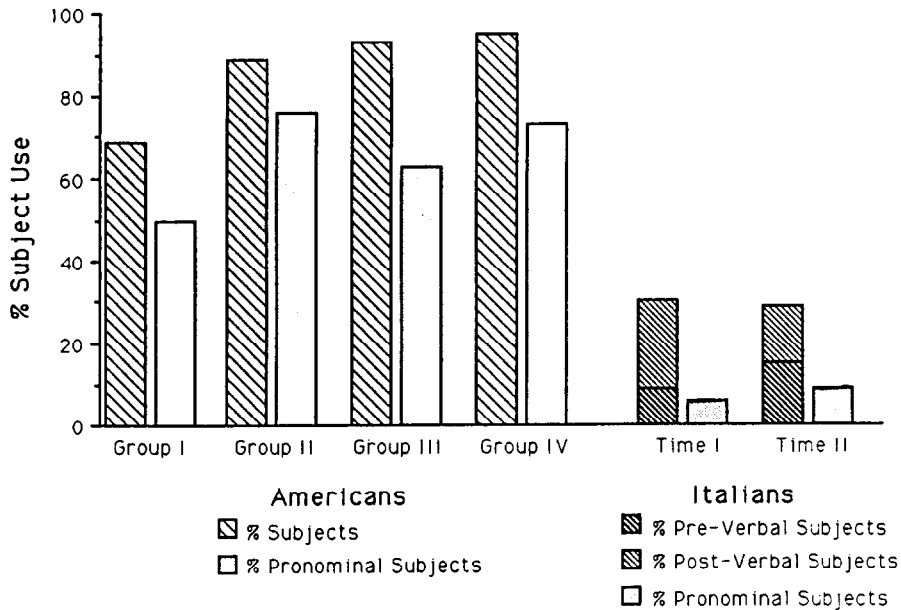


Table 3. *American children: close-up of Groups I and II. Proportion of utterances containing a verb which contain a subject noun phrase*

	MLU	Age	Proportion of subject use	Frequency of subject use ^a
Group I	1.53	2;1	.55	26/47
	1.74	1;10	.76	129/169
	1.79	2;2	.74	56/76
	1.81	1;10	.58	36/62
	1.99	2;2	.82	62/76
Mean	1.77	2;0	.69 (s.d. = .12)	
Group II	2.24	2;3	.90	95/106
	2.28	2;3	.84	191/227
	2.52	2;7	.85	234/274
	2.66	2;8	.91	166/182
	2.76	2;5	.94	236/250
Mean	2.49	2;5	.89 (s.d. = .04)	

^aThe denominator is the number of non-imitative non-imperative strings containing a verb; the numerator is the number of those strings containing a subject.

For 18 of the remaining 19 children there was either no change or a decline of one percentage point.

Pronominal subject use. Whether Measure 1 or Measure 2 is used to assess pronoun use, at every MLU group a majority of the children's subjects were pronouns. Further, for every child pronouns are more than half their subjects (the lowest figure was 59%), and usually more than three-quarters. Measure 1 includes *wh*-questions and classifies interrogative *wh*-words as pronominal subjects. On that measure, .77 (s.d. = .15) of Group I's subjects are pronouns, .87 (s.d. = .06) of Group II's, .85 (s.d. = .05) of Group III's, and .84 (s.d. = .04) of Group IV's. Had the true (adult underlying) subjects of *wh*-questions been scored as the subject, instead of the *wh*-words, the figures would have been almost identical, since in almost all cases the true subject was "this", "that", "these", "those", or "it". The results are shown in Figure 3 to the right of each bar representing subject use. Pronoun use is represented as the percentage of +V utterances containing a pronominal subject (Group I = 53; Group II = 77; Group III = 79; Group IV = 80).

Measure 2 excludes *wh*-questions and includes only personal and demonstrative pronominal subjects. Here the averages are slightly lower. For Group I there is a decline of 4 percentage points: 73% of all subjects were pronouns. For the remaining three groups the decline is only 1 point.

Parental use of subjects. In order to determine what the input to the child is, we calculated the parents' use of subjects in their immediate replies to children's utterances. We included only non-discards, non-imitative, and non-imperative utterances with verbs. Parents are very consistent in using subjects. In Group I, .96 (s.d. = .04) of usable parental replies included a subject, in Groups II, III, and IV, .98 (s.d. per group always below .02). When imperatives are included, the percentage of replies including a subject drops to .90 at Group I, and to .93 for Groups II, III, and IV.

Italian children

Since so few utterances per child were available for the Italian data, fewer analyses were performed.

Use of subjects and pronominal subjects. The total proportion of non-imperative non-imitative strings containing subjects, whether a verb was present or not, is .20 at Time I and .23 at Time II. Those figures are comparable to American children in Group I.

With respect to appearance of subjects in utterances containing a verb, Italian children look completely different from American children. The right-hand portion of Figure 3 shows the Italian children's use of subjects in utterances containing a verb. The top and bottom halves of the subject bars show pre-verbal and post-verbal subject use. From Time I to Time II the overall proportion of subject use stays constant at about .30 (average s.d. = .08), less than half the average usage of Group I Americans.

At Time I the children produce twice as many post-verbal (.21, s.d. = .04) as pre-verbal (.09, s.d. = .06) subjects. By Time II the children have almost doubled their production of pre-verbal subjects (to a proportion of .15, s.d. = .05), but total production of subjects remains the same from Time I to Time II. The Italian children differ from American children, in producing few subjects and in showing no increase in production within the time period sampled.

Compared to the American children's, the Italian children's total subject appearance is high relative to the use of subjects in utterances with verbs. In the American Group I data, the percentage of subjects in all strings was one-third the percentage of subjects in strings with verbs, whereas here it is two-thirds. In the Italian coder's opinion, many of the Italian strings lacking verbs consisted of a subject plus an object or adjective, or an adjective plus a postposed subject. The latter structure in American children's speech was considered to be simply a noun phrase.

The Italian children also produce less than half as many pronominal subjects as the American children. Because of the small number of utterances with subjects, the data were pooled rather than averaged over subjects. Pooling data resulted in slightly larger averages than averaging over subjects would have.

Figure 3 presents the proportions of subjects containing any sort of pronominal, whether it be a personal pronoun, a demonstrative, or an interrogative. At Time I 22% of the Italian children's subjects were pronouns, and at Time II 35% were pronouns. (For Group I of the American children, about 75% of the subjects were pronouns.) At both Times I and II, pronouns were more common in post-verbal than pre-verbal position.

Discussion

The data comparing American and Italian children show that American children look very different from Italian children with respect to use of subjects. Italian children look the way one would expect the learner of a prototypical Romance null subject language to look. From age 1;6 to 2;6 they include subjects in a minority of their sentences with verbs – about 30% – and they use pronoun NPs as a minority of their subjects – about 20–35%.

In classifying the Italian children's utterances we have assumed that the post-verbal subjects are genuine subjects that were postposed from pre-verbal position. On some analyses, however, the post-verbal subject is generated in situ, and *pro* is generated as the pre-verbal subject. If *pro* were considered to be the subject in all Italian sentences except those with a lexical pre-verbal subject, Italian children at Time I would be using subjects only 9% of the time, and at Time II only 15% of the time.

In contrast, American children in each of our MLU groups include subjects in most of their sentences with verbs. In Group I (our lowest-MLU group: MLU 1.53–1.99), almost 70% of the children's utterances with verbs include subjects – more than double the rate of Italian children. Even when *wh*-questions are excluded from the American data, the American children produce subjects at double the rate of Italian children. Group II children (MLU 2.25–2.76) use subjects consistently; the lowest percentage of subject use is 84% of utterances with verbs – still a high level of performance.

Further, Group I uses pronoun NPs for the majority of their subjects – over 70% – again more than double the Italian rate. We note a parallel in use of pronouns: just as American percentage of pronoun use increases, so does Italian. From Group I to Group II the American use of pronominal subjects increases from 73% to 86%; from Time I to Time II the Italian use increases from 22% to 35%. This suggests that in both groups of children their initial use of pronominal subjects is depressed for reasons independent of the status of subjects in the language.

How has the impression of low subject use among American 2-year-olds developed? Examination of Group II's total subject use indicates why subject use may appear to be low even in children who are using subjects consistently. The percentage of subjects in the whole corpus (and even these data do not include the

entire corpus, since they exclude discards, imitations, and imperatives) is low – 50%. Thus, in scanning a corpus, one can be misled by the large number of utterances that lack a verb into thinking that the child's use of subjects is low even when it is high.

If we take subject usage from 84% to 94% as evidence that children understand that subjects are obligatory, we can conclude that at least soon after MLU 2 American children exhibit no competence deficit. If the pro or VP hypotheses describe a stage in acquisition, it would have to be a stage before Group II.

What about Group I? Here we must pause. Two children (at MLUs 1.53 and 1.81) showed a low use of subjects: 55% and 58% subjects. When *wh*-questions were excluded, the lowest-MLU child's usage was only 38%. The other three children in the group averaged 77% (74% when *wh*-questions were excluded). Group I may consist of two subgroups, for one of which either the pro or VP hypothesis holds. Under MLU 2.0, then, some children may exhibit a competence deficit while others do not. With our sample we will be unable to resolve this question. Our procedure will be to continue to treat Group I as a group, while keeping in mind that the lowest-MLU child, in particular, may not conform to the rest of the group.

Another caveat concerning our data is that Bloom, Miller, and Hood's (1975) investigation of 4 children appears to show less subject use for 3 of their 4 children than our data would predict. Their data cannot be directly compared with ours, since they excluded most intransitive verbs. Nevertheless, the comparison of Bloom et al.'s data with ours suggests that there may be developmental differences in children's use of subjects. Only a larger database will answer this question.

We can now consider how well the competence deficit and performance theories account for Group I's data, and for the differences between the American and Italian children. What do the theories predict? Are they confirmed or disconfirmed?

If the competence deficit theories are interpreted as allowing no role for input effects or performance constraints, they would then predict equal usage by American and Italian children. On that interpretation the theories are disconfirmed, because the data show unequal usage. If the theories allow input to influence the child, and assume no performance constraints, they would then predict that Italian children's output would match Italian parents' and American children's output would match American parents'. On that interpretation also the theories are disconfirmed, since neither Italian nor American children produce subjects at the same rate that adults do. (Bates, 1976, supplies figures for two Italian parents in a total of three samples: the parents used subjects about 50% of the time. (Bates states the percentages as 30–40%, but my calculations from her data yield 46–56%.))

Only if the theories include a role for input effects and performance constraints

would they be confirmed by the data. If children try to match their input regardless of their grammar, and if they are limited in how fully they can match the input, then American children could have pro or an incomplete grammar, but be producing subjects a majority of the time and at a higher rate than Italian children.

Thus, the competence deficit theories can account for Group I's data and for the differences between the American and Italian children, but *only* if they add the ancillary assumptions about input and performance. Neither competence theory by itself can make any predictions about how the child's grammar will be manifested in the child's output.

The performance account directly includes a role for performance constraints, claims the child's knowledge is correct, and is neutral about the role of input. On that basis the performance account predicts that American children will use subjects less than 100% of the time, and predicts they will use subjects more than Italian children do. The data thus confirm the performance account with no fancy footwork needed.

On balance, then, our provisional conclusion is that American children know early on that their language requires subjects: neither the pro hypothesis nor the VP hypothesis consistently describes children above MLU 1.5. Since two of the Group I children show relatively low usage, however, and since ancillary assumptions could bolster the competence deficit hypotheses, our conclusion is tentative. The following studies look at other aspects of the children's grammars in an attempt to determine the soundness of our provisional conclusion.

Summary

American and Italian children at roughly comparable ages and levels of linguistic sophistication use subjects to different degrees. Group I Americans (MLUs between 1.53 and 1.99), while not entirely consistent in their use of subjects, use subjects and pronominal subjects about twice as much as Italian children. Neither of the two competence-deficit hypotheses explored would predict such a pattern, without being bolstered by ancillary assumptions. A performance hypothesis straightforwardly accounts for the data.

Study 2: Expletives and pronoun types

While sheer amount of pronoun use by the American children suggests that they are, at least, not operating with the *wrong* (Italian) value of the null subject parameter, the types of pronouns children use might help clarify matters in two ways. The first way concerns whether the children use expletive pronoun subjects, such as *it*. The second way concerns whether the subject pronouns the children use receive nominative case.

If the children think their language is a null subject language, and if being a null subject language entails lack of expletives, one might predict, as Hyams (1986) did, that children's speech would correspondingly lack expletive subjects. Hyams has claimed that American children only begin producing expletive subjects once they use subjects consistently. We therefore searched for the existence of expletive subjects in the American children's speech, and for "expletive" contexts in the Italian children's speech. An expletive context was defined as one in which a subject was obligatorily absent, as with the verb *piovere* – "to rain".

On the VP hypothesis children's grammars lack elements of the INFL system. (In some recent treatments, e.g., Pollock, 1989, INFL is not a node; instead, tense and agreement each have their own separate node. We shall retain the more usual conflation of both features under a single INFL node.) The feature tense or Agr, within INFL, assigns nominative case to subjects. If we add no input assumptions, it follows that if children lack tense, or an INFL system generally, they will be unable to reliably mark the case of subject pronouns. The case of pronouns should be random. One way, then, that we can determine whether the child's grammar is lacking INFL is to determine whether subject pronouns reliably receive nominative case. Accordingly, we established which pronouns were used as subjects, and what case they appeared in.

Method

The corpora from Study 1 were used. Utterances labeled discards were eliminated, as were imitative and imperative utterances.

Expletive subjects

American transcripts were searched, using a custom search program, for all occurrences of *it*, as well as for likely expletive contexts. The expletive contexts searched were *rain*, *snow*, *hot*, *cold*, *dark*, and *seem*. The Italian corpora were searched for verbs which require the absence of subjects: *bisognare* ("to need"), *piovere* ("to rain"), *bastare* ("to be enough"), *accadere* ("to occur"), *avvenire* ("to happen"), *sembrare* ("to seem"), *nevicare* ("to snow"), *fare freddo* ("to be cold" – weather), *fare caldo*, *fare notte*, *fare giorno*, etc.

Pronominal subjects

We counted the personal pronouns the American children used as subjects via a computer-assisted search. Among sentences with subjects, the following pronouns were searched for: *I*, *me*, *you*, *he*, *she*, *it*, *him*, *her*, *we*, *us*, *they*, *them*, *em*, *my*, *your*, *his*, *hers*, *their*, *theirs*. (For *you* and *it* case cannot be assessed.) We also manually searched for pronominal objects of verbs and prepositions, to determine whether children's nominative case-marking was vacuously correct. If children use the nominative form regardless of the position of the pronoun, they

cannot be described as knowing that the nominative case should be restricted to subjects. The Italian corpora were not searched for pronouns because the database was too small.

Results

Expletive subjects

For the American children, expletive use is infrequent at all MLUs. To the extent that there is use of expletives, it occurs across the board. Among the 21 children there were only 12 candidate instances of expletive *it*:

Child MLU	1.53	"(When it's noe-y)"
	1.79	"When it rains"; "It rains"
	2.76	"It's dark"; "It's dark outside"
	3.07	"It's winter time?"
	3.15	"It looks summertime, 'cause it's summertime"
	3.31	" 'Cause it's dark"; "When it's hot"
	4.12	"Okay, it's very cold"; "It's hot"

There were no instances of the expletive contexts that we searched that lacked an expletive subject. Thus, as far as can be determined, there were no instances where a context required an expletive and the child failed to provide one. Undoubtedly such instances occur in child speech, as they do in adult speech. But since the base rate of expletive contexts is low, it will be difficult to find examples in only one and one-half hours of taping. Given how infrequently high-MLU children, who are producing subjects more than 90% of the time, produce expletives, the low rate of production by low-MLU children is not meaningful. Expletives are not a good diagnostic for our data.

An incidental fact can be noted: 5 of the 12 expletives were in subordinate clauses. Considering how few subordinate clauses are produced, that percentage is very high.

The Italian children only produced examples of *bisognare* ("to need"), *piovere* ("to rain"), and *bastare* ("to be enough of"). At Time I there was one instance; at Time II there were a total of 14 instances, produced by 4 of the 5 children. All of the examples correctly lacked a subject. While the rate of production of expletive contexts is low, it is much more frequent than the American children's production.

Pronominal subjects

All the American children produced personal pronouns as subjects; no child produced fewer than three types. The lowest-MLU child, however, produced a total of only 5 tokens, 19% of subjects. Further, Group I as a whole produced a smaller percentage of personal pronouns as subjects (52) than did Groups II

(78), III (67), or IV (71). Group I, then, and especially the lowest-MLU child, need further scrutiny.

With respect to the case of the pronouns used, Table 4 shows that the children overwhelmingly used the nominative case when they produced pronominal subjects. Group II produced the largest proportion of incorrectly cased pronominal subjects, at 2% of total pronominal subjects: Group II was the only group in which every child produced at least one incorrectly cased pronoun.

There were no examples of nominatively marked pronouns in object positions. (A possible exception was the utterance "Putting they, in their bed", which appeared to us to be a false start.) For all groups, when the children used a pronominal object which received overt case-marking (e.g., *me*, *him*, *her*, *us*, *them*), the form was correct. But there were few examples of markable pronouns in object positions in Groups I and II: most pronouns used after a verb were *it*, *you*, *this*, or *that*. In Groups I and II only 25% of the pronouns were markable pronouns, contrasting with 80% object in subject position. In Groups III and IV objectively marked pronouns were common.

Discussion

When expletives are used, they are used by our children below MLU 2 as frequently as by our children above MLU 4. But expletives are rarely used by American children across the MLU range we observed. Even when children are producing subjects for 90% of their utterances with verbs, they appear to have little reason to utter expletives, doing so only 12 times in our corpora. The children do not use verbs, such as *seem*, which take expletive subjects, and they probably only speak about the weather when it is relevant to undertaking some activity. Contrary to Hyams' (1986) claims, then, we see no relationship between expletive use and subject use.

Table 4. American children: proportion of different types of pronominal subjects

Group	<i>I</i>	<i>you</i>	<i>he</i> <i>she</i>	<i>it</i>	<i>we</i>	<i>they</i>	<i>my</i> <i>me</i>	<i>him</i> <i>his</i> <i>her</i>	<i>em</i>	All Nom	All *	<i>N</i>
I	.63	.05	.15	.13	.02	.02	.003	0	0	.82	.003	34
II	.66	.06	.07	.12	.02	.04	.01	.01	.004	.80	.02	141
III	.49	.17	.14	.12	.03	.05	0	.003	0	.71	.003	174
IV	.44	.10	.16	.11	.13	.06	0	.002	0	.79	.002	169

Note. The final column is the mean number of personal pronouns used per child. "All Nom" is the proportion of nominative pronouns (*I*, *he/she*, *we*, *they*); "all *" is the proportion of incorrectly cased pronouns (*my*, *me*, *him*, *his*, *her*, *em*). The remaining pronouns (*you* and *it*) are not overtly case-marked.

The data from the Italian children support the interpretation that American children use few expletives for semantic rather than syntactic reasons. Although the Italian children produced expletive contexts more often than the American children did, their use appears to reflect the meanings of the verbs involved rather than the grammars of the two target languages. Of the three "expletive" verbs the Italian children used, two are clearly of importance to children: *bisognare* has to do with there being a need for something, and *bastare* with there being enough of something.

Certainly there is nothing in our data to suggest that expletive use is related to subject use. If expletive use is taken to be diagnostic of knowledge that subjects are obligatory, then even Group I children understand that. Their inconsistent usage of subjects could not be explained by the pro hypothesis.

The predictions of the VP hypothesis, which would explain the inconsistent subject usage by our lowest-MLU group as due to a grammar lacking INFL, are not borne out by our data. Contrary to prediction, even Group I children show uniform use of nominatively cased NPs in subject position. Nevertheless, there are two countervailing factors.

First, the lowest-MLU child produced particularly few personal pronouns as subjects. This was also the child who produced the smallest percentage of subjects. This child, then, may be best described in terms of the VP hypothesis, and the period between MLU 1.5 and 2.0 may be best viewed as a transitional period when some children have a full grammar and others have only a VP grammar. The remaining children in Group I showed usage of personal pronouns slightly below the higher-MLU groups, but within their range.

Second, the clearest distributional pattern would be uniform use of nominative pronouns in subject position and of objective pronouns in object position, as was the case for Groups III and IV. But Groups I and II behaved differently. Those children consistently used nominative pronouns in subject position but used few markable pronouns in object positions (using primarily *you*, *it*, and demonstratives). Groups I and II appeared to "avoid" using markable pronouns as objects, although they correctly case-marked the markable pronouns they did use in object position.

The most reasonable interpretation of Groups I and II's pattern (excluding the child at MLU 1.53) is that the children primarily knew the nominative forms of the pronouns they used, and also knew that the forms *were* nominative. Otherwise they should have used the nominative forms freely in all object positions, which they did not do. If the only way of establishing nominative case is via INFL, then the children also have INFL. If that is so, then the VP hypothesis cannot account for Group I's inconsistent use of subjects, and a performance analysis is correspondingly strengthened.

Whether the children assigned nominative case via INFL, however, or simply understood that the nominative form appears in subject position, cannot be

answered by our data. Input may play a role: the child might restrict herself to the pronominal forms that she has heard in subject position, which will generally be nominative. No change in the child's output would then be observed when case was assigned via INFL. The effect of adding this assumption is to rob the VP hypothesis of predictive force, since it will now predict both random use of pronouns and consistent use of nominative pronouns. A different ancillary assumption could be derived from Lebeaux (1987), who proposes that case is first assigned via the phrase structure configuration, and then via tense. Again, however, predictive force regarding case would be lost.

The lack of objectively or possessively cased pronouns is surprising, given cited examples of both (e.g., Budwig, 1985, 1989; Radford, 1990). At present we cannot say why such discrepancies exist. It may be that, although errors of using objective case in nominative position exist, they are generally infrequent in each child's usage relative to the number of correct uses, or that there is wide variation among children in their use of non-nominative forms.

Previous data summaries do not provide the comparative figures required to determine how common errors of incorrect case are. Data on first-person subjects, provided by Budwig (personal communication), indicate that, of the six children she observed, one used "me" for subjects about 25% of the time, while the rest showed little or no usage. Three frequently used "my" as first-person subject, with percentages ranging from 21% to 44%. Thus, there may well be children who show extensive usage of non-nominative forms in subject position. We need data from large numbers of children, so that developmental differences in acquisition can be identified and evaluated. Reanalysis of others' transcripts and data is likely to underestimate the true variation in acquisition. Our data strongly suggest that, from MLU 1.5 on, there is consistently correct use of subject pronouns.

Summary

There is no support in our data for either competence-deficit hypothesis as an explanation for Group I's behavior. Group I's inconsistent subject usage does not co-occur with other properties that would be expected if either competence-deficit hypothesis were correct. The pro hypothesis is not confirmed by the children's use of expletives: to the extent that children use expletive pronouns, they do so across a broad MLU range and subject use range. If the expletive data are meaningful, either our lowest-MLU children already understand that subjects are required, or expletives are irrelevant to the status of subjects. The VP hypothesis is also not confirmed, because the children appear to have correctly cased nominative pronouns in subject position. If the only way nominative pronouns can arise is via marking from tense or Agr, then Group I children, by implication, have INFL, or at least a grammar larger than VP.

Study 3: Modals

Thus far, the most straightforward predictions of competence-deficit theories have not been borne out by the data. American children show inconsistent use of subjects only at MLUs below 2, and even then they use subjects and pronominal subjects much more than Italian children do. They also use expletive subjects very early in acquisition, and have correctly case-marked pronominal subjects. Nevertheless, the children's early usage is limited. Children's use of modals potentially provides another way of assessing the significance of the subjects they use, since most competence-deficit theories have tried to explain a presumed absence of modals in young children's speech. It is therefore germane to measure children's use of modals, first to determine what the facts are, and second to assess the hypotheses about children's early grammars.

In English, modals originate under the INFL node, rather than under the VP node (see, e.g., Pollock, 1989). Modals do not have an infinitival form (* "to can"), nor do they inflect for person and number (* "she cans"). Modals in English do carry tense, and exist in present (*can, will, may*) and past (*could, would, might*) tense forms. In Italian (and French) modals are main verbs and originate under the VP node. In Italian (and French), there is an infinitival form, and inflections for person and number.

The pro and VP hypotheses appear to be on their weakest ground with respect to modals. While both have predicted absence of modals, there is little theoretical basis for the prediction. In Hyams' (1986) formulation of the pro hypothesis children projecting a null subject language had an element in INFL that simultaneously licensed pro subjects and ruled out the co-presence of a modal. Thus, an American child with the null subject parameter set incorrectly would also have the INFL element that prevented classification of modals as part of INFL. Hyams went on to argue that the children would also be unable to treat modals as main verbs, because of their lack of inflection. Having no place to put the modals, as it were, the children would fail to produce them altogether. Once their grammar changed so that the element in INFL licensing null subjects was absent, modals could then be classified.

As Lebeaux (1987) points out, however, the reasoning does not go through. Nothing in principle stops the child from initially treating modals as defective main verbs within VP. Indeed, if the child has the null subject value of the parameter, that should *force* her to classify modals as main verbs.

To rule out modals as main verbs, one would have to introduce an ancillary assumption stating that a verbal element must behave like other main verbs in all respects for the child to classify it as a main verb. Modals do behave like main verbs in *some* respects (e.g., they carry tense, though we do not know if the child recognizes this). The assumption therefore has to be that the child fails to classify modals as main verbs because they do not act like main verbs in *all* respects.

Further, the assumption would also have to apply generally to all verbs, requiring a child to see each verb in all its forms before classifying it as a verb. Since the assumption is not only unmotivated syntactically, but is also unlikely as a principle of acquisition, there are no good grounds for predicting absence of modals, only grounds for ruling them out as part of INFL.

As Lebeaux (1987) also points out, there is a more serious problem with the pro hypothesis. Say that the child *is* unable to classify modals as main verbs and say that the incorrect setting of the null subject parameter blocks the analysis of modals as part of INFL. If modals are recognized as having any verbal character at all, then they should force a change in the grammar, and, given their frequency in adult speech, do so reasonably quickly. In other words, on the logic of Hyams' (1986) analysis, modals should act as a telltale and cause the child to reset the parameter. (While Hyams footnotes a similar possibility, she does not recognize it as a problem.)

The situation is somewhat different for the VP hypothesis, since it is often accompanied by a maturation mechanism (Guilfoyle & Noonan, 1989). Rather than the input forcing a resetting of the parameter, the child's grammar changes as a function of maturation (see Borer & Wexler, 1987, for discussion of maturation). With such a mechanism the input is not a causal agent in grammar change. Thus, while the VP hypothesis alone does not allow a prediction that the child will ignore modals, the maturation hypothesis does (and in fact does so whether associated with the VP or the pro hypothesis).

The argument would take the following form for the VP hypothesis. The child has only a VP grammar; she cannot classify modals as verbs because of their inadequate verbal morphology (this reasoning has already been questioned); she cannot classify modals as modals because her grammar lacks INFL – therefore, no modals. Once the grammar matures to include INFL, she can recognize modals and include them in her output. The argument could be used for the pro hypothesis as well. The child has pro; she cannot classify modals as verbs; she cannot classify modals as modals because her grammar will not allow them as part of INFL – therefore, no modals. Once the grammar matures to allow modals as part of INFL, the child can recognize them and include them in her output.

The similar nature of the two arguments shows that the hypothesized form of the child's grammar is insufficient to predict absence of modals. The critical work is done by: (a) the assumption that the child cannot classify modals as verbs, and (b) the maturation mechanism. The prediction follows mainly from the claim that at Time 1 the child's grammar is too immature to allow correct classification, and at Time 2 it is mature enough to allow correct classification. In this case, at least, appeal to a maturation mechanism is an ad hoc and circular answer to the question of how children's grammars change.

Given the weak theory-internal basis for predicting absence of modals, it is reasonable to ask why absence *has* been predicted. The principal reason seems

to be the presumption, drawn from previous summary data (e.g., Bellugi, 1967), that modals are absent, even when semi-auxiliaries (such as *wanna*, *gonna*, and *hafta*) are present. Here we measure the use of modals and semi-auxiliaries to determine whether there is an absence to be accounted for.

Method

The corpora from Study 1 were used. Utterances labeled discards were eliminated, as were imitative and imperative utterances.

Measurement of modals

Using computer-assisted search procedures, we counted the number and types of modals the children produced. We searched for *can*, *can't*, *could*, *couldn't*, *'ll*, *will*, *'ll*, *won't*, *would*, *wouldn't*, *shall*, *should*, *shouldn't*, *may*, *might*, *must*, and *ought*. We searched the Italian corpora for all forms of *dovere* ("must") and *potere* ("can").

Measurement of semi-auxiliaries

We searched for *gonna*, *wanna*, *hafta*, and *gotta*. We also searched for and included in our tabulations *going to*, *want(ed) to*, *have to*, *had to*, and *get to*, even though those forms could also be considered as full verbs with infinitival *to*. We did not search the Italian corpora for such verbs.

Data tabulation

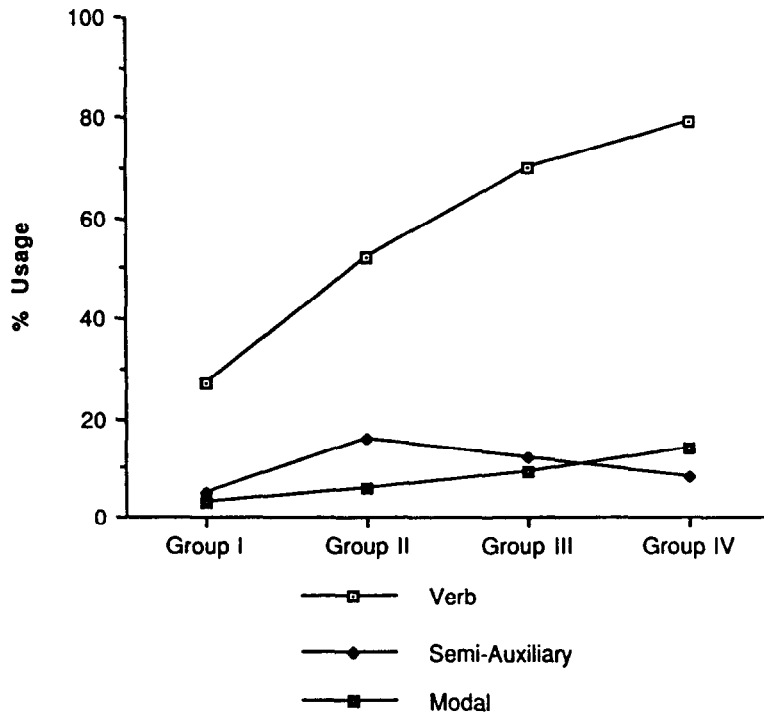
To calculate the proportion of modals and semi-auxiliaries we used the number of utterances containing a verb as the denominator. (Some utterances contained more than one verb; such utterances were only counted once.) The numerator consisted of the number of modals or semi-auxiliaries (seldom more than one per utterance).

Results

All American children except the lowest-MLU child produced modals. In every group, *can* and *'ll/will* were the two most common modals. Figure 4 graphs the percentages of modals, semi-auxiliaries, and, for comparison, verbs, in each group. Modal usage was infrequent in Group I: only 3% (s.d. = 3) of utterances with verbs contained modals, a total of 14 tokens. There was a steady gradual increase: to 6% (s.d. = 2) in Group II (61 tokens), 9% (s.d. = 4) in Group III, and 14% (s.d. = 2) in Group IV.

Italian children produced even fewer modals than Group I Americans, at both Time I and Time II. At Time I there was only one modal: at Time II 1.5% of utterances with verbs contained modals. Since the status of the modals in the American children is the main interest, we will present those results in detail.

Figure 4. American children's production of verbs, semi-Aux's, and modals.



American modals

In Group I, *can*, *'ll*, and *will* were produced. In Group II, *can't*, *could*, *won't*, *would*, and *may* were added. In Group III, *couldn't*, *should*, *might*, and *must* were added. In Group IV, *shall* and *shouldn't* were added, but *might* and *must* did not appear. Fewer past tense modals were produced than present tense ones. The first past tense modal appeared in the child at MLU 2.52.

Contrary to what one might have expected from Bellugi's (1967) data, the earliest modals were not contracted with a negative. Not until the child at MLU 2.28 did a modal appear contracted with a negative; *can't* appeared with the child at MLU 2.28, and *won't* appeared with the child at 2.52. Only one child, at MLU 2.76, produced only modals contracted with a negative: *can't* and *won't*.

Modal usage increases gradually and steadily as a function of age and MLU across all four groups, and there is no hint of a step function. The Pearson product-moment correlation between MLU and modal usage was $.71$, $p < .001$, and between age and modal usage $r = .66$, $p < .001$ (two-tailed). The partial correlation between MLU and modal usage (with age partialled out) just missed significance, $r = .43$, $p = .056$. The partial correlation between age and modal

usage (with MLU partialled out) was low, $r = .28$, n.s. Thus, MLU and age jointly correlate strongly with modal usage, and MLU alone shows a stronger relation with modal usage than does age alone.

Modal usage neither suddenly begins when subject use becomes more consistent in Group II, nor dramatically increases when subject use becomes more consistent. To examine the relationship between modal usage and subject usage, we correlated the proportion of subjects in utterances with verbs (data shown in Figure 2) with modal usage. The simple correlation was high, $r = .58$, $p = .006$. However, when MLU and age were partialled out, the correlation dropped to $-.04$. Thus, modal usage has no independent relation with subject usage. Children's use of subjects is roughly constant regardless of how many modals they produce.

To investigate more closely the relation between subjects and modals we also calculated what proportion of utterances with modals included a subject. The data can only be suggestive in Group I, due to the small number of tokens. In Group I, the average proportion was .94; in Group II, .95; in Group III, .98; in Group IV, .99. Children use a higher proportion of subjects in utterances with a modal than they do on average in all their utterances with verbs (compare with Figure 3).

American semi-Aux's

The lowest-MLU child produced no semi-Aux's, but all other children produced them. The results are graphed in Figure 4. As with modals, production of semi-Aux's was infrequent in Group I: 5% (s.d. = 6) of utterances with verbs contained semi-Aux (19 tokens). As can be seen from Figure 4, usage of semi-Aux's shows a complex relation with age and MLU. Usage peaks during Group II (16%, s.d. = 8) and then gradually diminishes in Groups III (12%, s.d. = 8) and IV (8%, s.d. = 5). There was also much more variability of semi-Aux usage within each group than there was for modal usage. *Gonna* and *wanna* were the two most common semi-Aux's.

We calculated what proportion of utterances with semi-Aux's included a subject. Again, for Group I, the data are only suggestive. In Group I, the average proportion was .63; in Group II, .89; in Group III, .87; in Group IV, .95. The proportion of subjects in utterances with semi-Aux's roughly parallels the proportion of subjects in all utterances with main verbs, in contrast to the proportion of subjects in utterances with modals, which is higher than the proportion of subjects in all utterances with main verbs.

Italian modals

The Italian children produced fewer, rather than more, modals than did the Group I Americans. For the Italian children at Time I there was 1 example of a modal ("dopo poi?" – "later can (you)?"). At Time II there were 6 examples –

about 1.5% of the verbs (4 produced by 1 child: 1 each produced by 2 children, out of a total of 392 utterances with verbs). A possible seventh example (“devi mica” – “(you) must not”) was interpreted as an imperative and therefore not included.

Discussion

Contrary to the predictions and claims of competence-deficit models, there appears to be no direct or primary relationship between onset of use of modals and obligatory use of subjects. Since neither model had strong grounds for predicting the absence of modals, their presence in American children’s speech need embarrass neither. The hypotheses were used to explain a presumed fact: once the fact vanishes, there is nothing to explain.

The children use more subjects and more modals as MLU advances, but once MLU or age is partialled out there is no correlation between modals and subjects. Further, the American children’s usage is greater, not less, than Italian children’s usage, even though modals are main verbs in Italian. (It should be noted, however, that there are more modal forms in English – *can*, *will*, *shall*, and *may* – than in Italian – *potere* and *dovere*.) Finally, modals are not first apparent as carriers of negation, nor do semi-Aux’s appear before modals appear. The conclusions based on the three children Bellugi (1967) examined are not verified by a larger sample of children, and the differences Hyams (1986) claims between American and Italian children are not confirmed in our study.

Our claim that Group I children have the category modal hinges on whether the early uses are genuine uses, in which the words are classified as modals or main verbs, rather than simply being uncategorized words. Only if the child’s initial correct, albeit limited and sporadic, use of a form (e.g., modals) is genuine, do we have evidence that, say, the child has an INFL node.

We have adopted the criterion of any amount of correct use. Although it runs the risk of letting in purely formulaic uses, there are also considerations in favor of it. First appearance correlates highly with more stringent criteria (Stromswold, 1989), suggesting that the child’s initial correct uses are the genuine seeds of knowledge out of which further knowledge and performance develop. It is also extremely unlikely that the first appearance in a corpus is the first time the child herself has used the form. Our samples are a minute fraction of the child’s productions. Sampling considerations alone would suggest that a child has been using a form before our data demonstrate it. Further, acquisition of any form has to start somewhere. There is no reason to think that the initial genuine uses would be widespread and consistent; quite the contrary.

At present any criterion beyond initial correct use appears arbitrary. Since Group I children made no distributional errors with the modals they used, we have called their use genuine. Although genuine, however, the modals may have

been classified as main verbs by the children rather than modals. But there is distributional evidence that the children have the modals categorized correctly. First, they use a different proportion of subjects before modals than before verbs; second, they use a different proportion of subjects before modals than before semi-Aux's. Modals and semi-Aux's are similar, in that they both precede main verbs, but they are treated differently by the children. Semi-Aux's are treated as main verbs, but modals are not. O'Grady et al.'s (1989) data roughly confirm ours. Third, the children do not use an infinitive after a modal.

We can advance an explanation for the asymmetry in subject usage with modals and semi-Aux's, harking back to an issue raised in the general introduction. It is less acceptable in adult speech to begin an utterance with a modal than with a semi-Aux. While both are ungrammatical, "Will play" is less acceptable than "Gonna play". The children's input probably reflects that difference: a parent might ask a child, "Gonna play?", or "Wanna play?", but will never ask her, "Will play?" The children appear sensitive to that asymmetry.

One reason, then, that the Group I children might have fewer modals than other groups is that producing a modal requires them to produce a subject as well. In contrast, producing a semi-Aux does not require them to produce a subject. A different reason for the asymmetry is that the extended projection principle obligates the child to include a subject if a modal is used.

Although we cannot, from our data, unequivocally conclude that the children have correctly categorized modals as modals, we should note that there is no distributional evidence in favor of the main verb analysis. The forms look like modals, and with development there is no indication of reanalysis; rather, a steady progression is seen. Thus, we tentatively conclude that even Group I children have a rudimentary use of modals.

Summary

Both the data in Study 2, indicating basic, and correct, nominative case-marking, and the data from this study, indicating primitive but correct use of modals from early in Group I on, argue that the child has an INFL node before MLU 2. If that is so, then neither competence-deficit analysis of Group I's inconsistent use of subjects is supported. The children's inconsistent use of subjects would reflect neither the presence of *pro* in the grammar (since *pro* is incompatible with a modal in INFL on one analysis; Hyams, 1986), nor an immature grammar consisting simply of a VP.

At the same time, the lowest-MLU child has consistently lacked the elements that the other children have produced. The VP hypothesis thus remains a good candidate description of the child's first grammar, and allows us to suggest how competence and performance factors might interact. The first grammar is ideally suited to the child's production constraints: the grammar will only allow short

utterances and the child's limited production abilities only allow short utterances. The child's reach and grasp have the same span. After the child's grammar expands, however, the child's reach exceeds its grasp. The grammar allows for longer utterances than the child can consistently produce.

Study 4: *To*, tense, and subordinate clauses

INFL plays a central role both in the adult grammar and in language acquisition accounts. The principal elements in INFL are tense features and agreement features. Tense can be realized either as –tense, with infinitival *to*, or +tense, with the two possible values being present or past. INFL elements assign nominative case to subjects; the subjects of untensed clauses do not receive nominative case but objective case from the preceding verb (contrast, e.g., "I wanted her to go", with, *"I wanted she to go").

The two competence-deficit theories make opposite claims about the existence of INFL. The pro hypothesis predicts the presence of INFL, and the VP hypothesis the absence of INFL. Since pro requires a licenser, which is presumed to be a feature in INFL, INFL must be present if pro is. Since the VP hypothesis limits the child to lexical categories, INFL is required to be absent. The most direct prediction the pro hypothesis makes is therefore the presence of INFL features, such as tense or agreement, in child speech. The most direct prediction the VP hypothesis makes is the absence of those features.

In both cases, however, the predictions are only possible if there are no other factors affecting children's productions. But Lebeaux (1987) and Kazman (1988), for example, note that the child could have tense or agreement features as an affix on the verb, in the absence of an INFL node. Ancillary assumptions have to be appealed to in order to secure a prediction of absence or presence of an element in speech. Neither hypothesis can make predictions about the child's productions solely on the basis of the hypothesized formal representation. In the case of O'Grady et al.'s (1989) proposal concerning the relationship between subject use and tense, their prediction would be little or no use of tense in children who use subjects under 90% of the time.

Tense is also important in Roeper and Weissenborn's (1990) analysis. On their analysis there is one and only one characteristic of null subject languages: null subjects (*pro*) can exist in embedded tensed clauses; in non-null subject languages they cannot. The interesting language acquisition consequence of Roeper and Weissenborn's claim is that only one feature in the input can serve as a telltale to set the correct value: the presence of null subjects in tensed subordinate clauses. Roeper and Weissenborn claim empirical support for their position. For example, they state that French and German children continue to use null subjects even after they use tense productively and use expletives. Therefore, exple-

tives cannot be serving as a telltale for null subjects, nor can INFL be sufficient for setting the correct value.

With respect to the American child, Roeper and Weissenborn have predicted (1990) that obligatory subjects will appear once the child has subordinate clauses. Since, on their analysis, the telltale of a null subject language is that a null subject can appear in tensed subordinate clauses, Roeper and Weissenborn predict a linkage between the child's production of subordinating conjunctions and complementizers (such as *that*) and the child's consistent production of subjects in subordinate clauses and in matrix clauses.

Method

The corpora from Study 1 were used. Utterances labeled discards were eliminated, as were imitative and imperative utterances.

Infinitival to

We searched for *to*, and tabulated the number of infinitives. We did not include *wanna* or *gonna*, but we did include cases of *want to*, where the transcript thereby indicated that two separate words were used.

Tense and agreement

We used a custom-made frequency count program to list all the lexical items in each child's corpus. We manually examined each child's list for any words that could be a past tense verb or a verb marked for third person singular. We then searched each corpus for the candidate verbs, and tabulated the instances of past tense verbs and third person singular verbs.

Subordinate clauses

We searched each child's corpus for subordinating conjunctions which introduce a tensed clause: *that*, *what*, *because*, *who*, *where*, *when*, *why*, *how*, *if*, *so*, *for*, *after*, and *before*. We tabulated the instances of subordinating conjunctions in tensed clauses and calculated the percentage of subject use in that context.

Results

Infinitival to

The infinitive was used by every group, but infrequently in Groups I and II. When infinitival use is relativized to verb use, the percentage of infinitives by group is 1.7, 1.4, 4.6, and 5.6. *To* usage increased irregularly with group, and most markedly between Groups II and III. There was a great deal of variability in frequency of usage.

In Group I, there was a total of 7 *to*'s (produced by 3/5 children; range per

Table 5. *American children: mean token frequency and proportion use of past tense*

	Mean frequency per child			Mean total proportion ^a
	Regular	Irregular	Total	
Group I	2.6	4.4	7.0	.09
Group II	4.2	9.0	13.2	.06
Group III	5.3 ^b	13.6	18.9	.06
Group IV	12.7 ^b	10.7	23.3	.09

^aThe numerator is total number of past tense examples and the denominator is number of utterances with verbs.

^bThese figures include irregular verbs which were incorrectly given a regular past tense suffix.

child from 0 to 3). Even the child at MLU 1.53 produced two examples of *to*: "(Oranges *to*) eat" and "Right, trying *to* go through it". In Group II, there were 17 *to*'s (produced by 4/5 children; range from 0 to 8); in Group III, 92 (produced by 8/8 children; range from 1 to 41, with 65 being produced by 2 children); in Group IV, 40 (produced by 3/3 children; range from 4 to 24).

Tense and agreement

Every child in every group produced tokens of past tense. Range of usage varied from 2% to 22% of verbs (with no correlation with MLU). Two children produced no examples of regular (*-ed*) past tense: MLU 1.53 and MLU 3.72. All other children produced examples of both regular and irregular verbs in past tense. Table 5 shows that while the average *frequency* of past tense use increased with group, the *percentage* of verbs that were tensed did not. Since the percentage of utterances with verbs increases dramatically with group, it is necessary to use verb production as a baseline. (The relatively high figure of 9% for Group IV is due to one child who produced past tenses for 22% of her verbs; the other 2 children produced 4% and 2%.)

There were no errors of incorrect form of the past tense of irregular verbs (e.g., "runned", "maked") until Group III, where 3/8 children produced one such mistake, and 2/8 produced 3 each. Even here, then, the errors were infrequent. Only a total of 9 incorrect tokens were observed in Group III, compared to 109 correct tokens, a proportion of .08. In Group IV, 1/3 children produced tensing errors: that child persistently mistensed one verb 15 times.

All but one child (MLU 1.81) produced at least one example of third person present singular *-s*. Across all the children there was a single error of inappropriate use of *-s*: "I shows you a cookie" (MLU 2.76). As with past tense, there was an increase in frequency with group, but no increase in proportion until Group IV. In Group I the children produced an average of 3.8 *-s*'s, or 4.2% of their verbs.

Group II produced an average of 8.4, or 4.2% of their verbs. Group III produced an average of 13.3, or 4.4%, and group IV produced an average of 22.6, or 9%.

Subordinate clauses

No child ever failed to use a subject after a subordinating conjunction introducing a tensed clause, but there were few examples in Groups I and II. In Group I there were only 4 instances of a tensed subordinate clause (among 3/5 children, including the lowest-MLU child); in Group II there were 11, in Group III 77; in Group IV 40.

Discussion

In all groups the children produced examples of all forms, but Groups I and II produced few infinitivals and few subordinate clauses. (The lowest-MLU child produced two infinitives, no regular past tense forms, one third person singular form, and one subordinate clause.) When there were major increases in frequency of usage of *to*, tense, and agreement (relative to number of verbs used), they occurred either between Groups II and III, or between Groups III and IV. Without ancillary assumptions, neither competence-deficit model accounts for the pattern of the data.

The children produced a relatively stable percentage of past tense verbs (5–10% of all verb uses) and third person present singular verbs (4% of all verb uses) over Groups I–III. There were few examples of past tense and agreement in Group I, but considering how few verbs those children used, they would have had to produce a very high percentage of tense and agreement to bring their frequency figures up. The increase of use of *to* between Groups II and III corresponds to findings by Bloom, Tackeff, and Lahey (1984) showing a similar increase in *to* somewhere between MLU 3.0 and 3.5.

As with subject use, it is clear how an impression of lack of tense, agreement, and infinitives could be gained from perusing young children's corpora. There *are* few examples in low-MLU children, and many more examples with higher-MLU children. In the case of past tense, however, there is no proportional increase in our data, once use of tense is relativized to verb use. O'Grady et al. (1989) state that the three children they observed produced almost no examples of tense until they began using subjects 90% of the time. The statement is slightly misleading, however, since their figures do show inflected verb use for all three children, and many examples for one child. Their tabulated data appear to agree with ours, though their conclusion differs.

We can briefly consider whether the children's early uses of *to*, tense, and agreement are genuine. With respect to tense, we note that errors of creating a regular past tense for an irregular form did not occur until Group III. Even in Group III there were very few examples – about 9% of the irregular verbs; errors

of overgeneralization are always fewer than correct uses. Marcus et al. (1990), analyzing data from a child database, similarly report low rates of overgeneralization over a wide age range, including ages comparable to children in this study.

It is often claimed that pre-overgeneralization children have simply memorized the past tense forms, and have performed no analysis that identifies the verbs as past tense (see Marcus et al., 1990, for discussion). That claim seems untenable unless there are specific errors, as have been attested for some children for some verbs (Kuczaj, 1981), of actively treating the past tense forms as untensed bare forms, by, for example, affixing a third person singular *s*, or *ing*, to the past tense verbs. Such errors appear to be rare and sporadic, however, and do not occur in our corpus. We favor the analysis of Marcus et al. that children's "overgeneralization" errors are actually retrieval errors, and consider the children's past tense uses imperfect but genuine.

With respect to agreement, the children's uses also seem genuine. Only one error of using the *s* ending inappropriately was found, and it is difficult to see how the child could restrict use of *s* to third person present singular without having coded the restrictions. That need not mean that the agreement feature is present in INFL, only that it is analyzed as a labeled agreement affix on the verb. Both tense and agreement, then, could be analyzed as features on the verb. We note that Meisel (1990) has independently concluded, in a study of three bilingual French-German children, that children show evidence of INFL before MLU 2.

To some extent the issue of genuineness is orthogonal to an evaluation of the two competence-deficit explanations for Group I children's inconsistent subject usage. Assume first that the early usage of *to* and tense in Groups I and II is not genuine, and that only in Group III, when there occurs a large increase in use of infinitives, the first errors of over-regularization of the past tense, and widespread use of subordinate clauses, is INFL well documented. That would be a problem for both hypotheses.

The problem for the VP hypothesis is that those changes occur at the wrong point. Instead of occurring between Groups I and II, which is when the major change in frequency of subject usage occurs, they occur between Groups II and III. The VP hypothesis would be forced to claim that the consistent subject usage in Group II still reflected a VP-only grammar, and would have no explanation for the large increase in subject usage from Group I to Group II.

For the pro hypothesis, which, on my interpretation, requires an INFL node to license a pro subject, there is a different problem. If there is no INFL until Group III, then there is no basis for a pro subject until Group III. In that case the inconsistent subject usage in Group I must have a different source than a pro subject, and the shift in usage between Groups I and II could not be explained by the pro hypothesis.

Assume now that the early uses are genuine (with the possible exception of the child at MLU 1.53). The VP hypothesis still cannot explain Group I subject

inconsistency, because INFL will already have been in existence, and therefore something else – such as production limitations – will be accounting for subject inconsistency. The pro hypothesis might be thought to fare better. If INFL exists even in Group I, then a pro subject is possible for Group I children. We would, however, still be left with the absence of a development between Groups I and II that would indicate a restructuring of the grammar.

Thus, the timing is off for both competence-deficit hypotheses. Whether we suppose the early uses of *to* and tense are spurious or genuine, we do not have the relationship between those elements and use of subjects that either hypothesis would require.

The data on the production of subordinate clauses were intended to bear on Roeper and Weissenborn's (1990) analysis. Our data suggest either that Group I children already have knowledge of subordinate clauses, or that knowledge is delayed until Group III. If the former, then Group I children should be taken to understand that subjects are obligatory; their inconsistent production of subjects would require another explanation. If the latter, then the Group II children are a problem, because they appear to have an obligatory subject, but produce few subordinate clauses.

If the early subordinate clauses are taken to be genuine, and if the children in Group I are taken to understand that subjects are obligatory, there is some support for Roeper and Weissenborn's (1990) prediction that subordinate clause use will be linked with subject use. If, however, subordinate clauses are interpreted as not developing until Group III, when they dramatically increase in frequency, then Roeper and Weissenborn's prediction is disconfirmed, since the children in Group II show good evidence of understanding that subjects are obligatory.

How should we understand the low production of subordinate clauses in Groups I and II? Subordinate clauses both require knowledge of individual subordinating conjunctions, and, unless the clause is used alone (as it was for the child at MLU 1.53), of embedding. Neither would be expected to be well developed at very low MLUs. A similar reason can be offered to explain why there are so few infinitives in Groups I and II (less than 2% of the utterances with verbs). Infinitives could increase because the higher-MLU children can handle the concomitant increase in sentence complexity and length that use of infinitives entails. We therefore favor the interpretation that the children's early uses are genuine.

Summary

If the children's uses of *to*, past tense, and third person singular present are considered to be genuine from the time of first appearance, the children then must be credited with having INFL at least in the middle of Group I. If the uses

are not considered genuine until they are more frequent, then the children do not show clear presence of INFL until after they are already using subjects consistently. In either case, both competence-deficit hypotheses fail to account for the facts.

We appealed very broadly to considerations of complexity and length as an explanation of increase in usage of infinitives and subordinate clauses, but in so doing we did not account for the precise timing of the increases. The next study spotlights performance issues more directly.

Study 5: Verbs, direct objects, and subjects

Thus far we have found no support for the pro hypothesis. There is some slight support for the VP hypothesis in the data of the lowest-MLU child, who produced the smallest number of subjects, the smallest percentage of nominatively cased subjects, no modals or semi-Aux's, no regular past tense, only two *to*'s and one subordinate clause. While one child's data are only suggestive, there is at least one child whom the VP hypothesis appears to describe reasonably accurately.

If that child does have a VP grammar, then the other children in Group I, who look importantly different from that child, have a grammar larger than a VP. In that case, they either have some other, unknown, competence deficit, or there are performance factors that constrain how often they produce subjects and other sentence elements. In other words, if we accept the lowest-MLU child as having a VP grammar, we must accept the others as having a fuller grammar, and correspondingly search for a different source for those children's subject omissions.

In Study 5 we look more directly at performance measures: does VP length vary as a function of type of subject; is subject use related to age, MLU, or verb use; are subjects and objects omitted equally often?

Performance accounts predict that the longer the VP the less likely a lexical subject will be chosen. L. Bloom (1970) and P. Bloom (1990) have proposed that the longer the VP, the more cognitive load is imposed, and therefore the greater the likelihood of dropping a subject. VP length was analyzed by L. Bloom (1970) for one child's utterances with the verb *make*, and by P. Bloom (1990) for three children's utterances with past tense verbs and verbs that cannot be used in imperatives (to exclude ambiguous utterances). In both cases, longer VPs were associated with subject absence. In addition, P. Bloom found that VP length was shortest when the child used a full lexical subject, longer with a pronominal subject, and longest with no subject. Since in both studies the sample size was small, and since past tense forms are, as we have seen, a minority of children's productions, it would be desirable to replicate those findings.

(Bloom, Miller, and Hood (1975) demonstrated that many factors affect chil-

dren's utterance length. Unfortunately, the large number of coding differences between their study and the present one precludes direct comparisons. Further research, however, should be aimed at systematically examining the effects of discourse and different types of complexity on subject use in particular. Here we restrict ourselves to a small subset of possible performance predictions.)

Performance accounts also predict that subject use will increase regularly as age and MLU increase. Since MLU is a measure of utterance length, it is obvious that increasing subject usage will increase MLU. But MLU could also increase independently of subject use, since increased use of any sentence element (e.g., complement clauses) will increase MLU. Age, though an imperfect measure of performance capability, is independent of any of the language variables. If a positive relation between age and subject use is found throughout the age range, that would indicate a performance component.

Performance accounts thirdly predict a relation between verb usage and subject usage. Verb production is operationally independent of subject use as measured here. The children could maintain a constant proportion of subject usage as their verb production goes up. But verbs introduce structural complexity into an utterance, especially if the child understands that utterances with verbs require subjects. If children's omission of subjects is related to a general limitation on the number of major sentence constituents that can be included, then one would expect subject usage and verb usage to be highly correlated. As children become able to handle the complexity involved in including verbs, they should correspondingly become able to handle the complexity involved in including subjects. That development should hold for both American and Italian children, since both should experience production limitations. However, it should be more extreme in American children, because the American range of subject use can be much greater.

Performance accounts might also predict asymmetries between use of subjects and objects. Pragmatic factors would lead to asymmetries because the given, and hence dispensable, information will tend to be the subject, and the new, and hence important, information will tend to be the object. Production factors could also be important, since the beginning of an utterance will be more effortful than the end. Finally, the utterance-initial location appears prone to omission and reduction effects, perhaps for prosodic reasons (Gerken, 1991).

Competence theories have been inconsistent in their predictions about asymmetries between subject use and object use. One theory (Kazman, 1988) has directly predicted absence of objects as well as absence of subjects. Hyams (1987), in contrast, explicitly claimed an asymmetry between subjects and objects. Radford (1990) has argued that both subjects and objects in fact are absent in early child speech, since the child freely allows "implicit" arguments. Thus, there is no uniform prediction from one competence-deficit account to another.

Finally, a performance account predicts two other developmental changes. The

first is an increase in the use of purely transitive verbs, verbs for which objects are obligatory: one way the beginning speaker can lighten the burden of producing objects for verbs is to produce more verbs that do not require objects. The second is an increase in how often children supply objects for "mixed" verbs – verbs that are grammatical with or without an object.

Method

The corpora from Study 1 were used. Utterances labeled discards were eliminated, as were imitative and imperative utterances.

Relation between type of subject and length of VP

We examined simple one-clause, one-verb sentences, excluding utterances with modals or negatives, and excluding imperatives and imitations. We computed the length in morphemes of verb phrases with a full lexical subject, a pronominal subject, or no subject; all verbs were used, including *be* as main verb. All material in the VP (complement or adjunct) was included.

Verbs and objects

We divided the children's verbs into three categories: those which we considered to be pure intransitives; those which we considered to be pure transitives; and those which could be transitive or intransitive. For Group I an average of 64 verb tokens per child contributed to the analysis; for Group II, 134; for Group III, 160; for Group IV, 168. We tabulated the proportion of verbs of each category as a function of MLU group.

We also tabulated the number of times direct objects appeared with the verbs of each category. Examples of utterances we classified as having an object with an intransitive verb are "he, he talking a duck", "go school", "yeah, it looks a lady", "I come teeth". *Be* was excluded from the tabulation; as a copula, it does not take objects in the same sense as the other verbs. The verbs *put* and *get* were also excluded; *put* requires two objects and thereby differs from the other verbs; *get* proved very difficult to score. Examples with an S complement or an adjunct phrase (e.g., "I go now") were excluded.

Results

Relation between type of subject and length of VP

The performance prediction that subject presence or absence would be related to VP length was supported, replicating L. Bloom (1970) and P. Bloom (1990). The predicted ordering is: longest VP with no subject, intermediate with a pronominal subject, and shortest with a full lexical subject. There are six possible orderings, so that the probability of any particular ordering appearing is .17.

For Group I, which would be expected to show the pattern most strongly, the average VP length was 2.53 with no subject, 2.45 with a pronominal subject, and 2.3 with a lexical subject, with 3 of the 5 subjects showing the phenomenon. For Group II, the figures were 3.08, 2.84, and 2.67, with, again, 3/5 children showing the phenomenon. Since there are six possible patterns, by chance 17% of the 10 children would show the predicted pattern; instead, 60% showed it. Application of the binominal test shows that the probability of 60% of the subjects showing such a pattern is less than .003. For Groups III and IV there was no consistent pattern. Only 3 out of the 11 children in those two groups showed the pattern which was dominant in Groups I and II.

Correlations among subject use, MLU, age, and verb use

Both predictions concerning correlations with subject use were borne out in our sample. The simple correlation between MLU and subject use is .77 ($p < .001$), and the simple correlation between age and subject use is .74 ($p < .001$). When a partial correlation between MLU and subject use is computed, the correlation is .48 ($p = .03$), a still significant, though much lower, correlation. The partial correlation between age and subject use is .41 ($p = .075$). MLU and age together predict subject use better than MLU or age alone.

Children's verb use increases markedly from Group I to Group IV, as Table 1 shows. Verb use is very highly correlated with MLU, even with age partialled out ($r = .81$, $p < .001$), but there is no correlation between verb use and age when MLU is partialled out ($r = .20$, n.s.). Verb use (unlike modal use) is highly correlated with subject use, even when MLU and age are partialled out. The partial correlation is .78 ($p < .001$). As children produce more utterances with verbs, they are correspondingly more likely to produce subjects for those verbs.

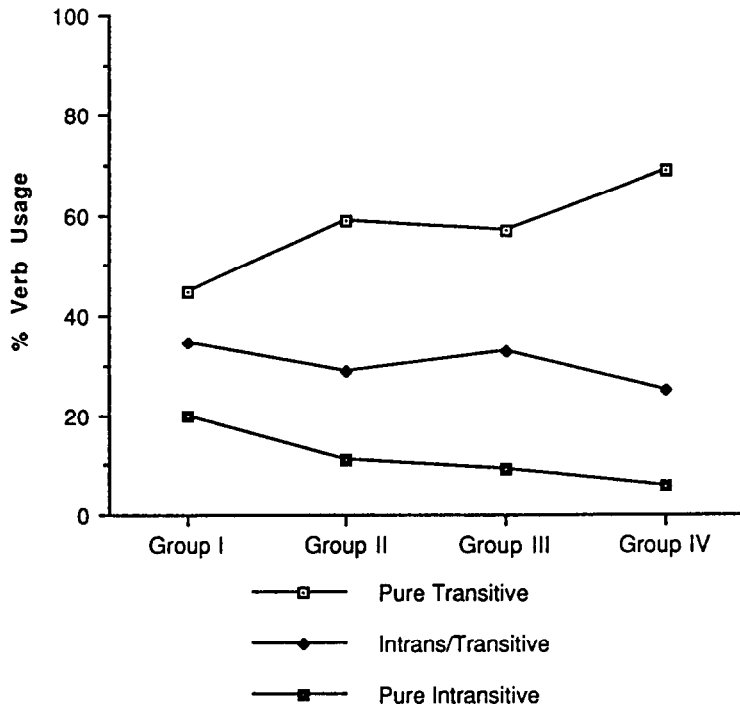
Because of the small number of Italian children, it is not possible to assess the relationship there between subject and verb use. Using a Spearman rank order correlation we find that at Time I the correlation is .50, and at Time II .90, but neither is significant at the .05 level (two-tailed). Nevertheless, it appears that there is a rough relation which is similar to that found in the American children.

Verbs and objects

As would be expected on a performance explanation, the children produced more purely transitive verbs as development proceeded. Figure 5 shows how the proportions of the children's verbs in the three major categories changed as a function of group, with the largest changes occurring between Groups I and II: the proportion of pure transitive verbs increases at that point, and the proportions of the other two verb categories decrease.

First, the proportion of pure transitive verbs increases from .45 (s.d. = .13) in Group I to .59 (s.d. = .10) in Group II. There is another increase from .57 (s.d. = .10) in Group III to .69 (s.d. = .11) in Group IV. Second, the proportion

Figure 5. American children's production of pure transitive, pure intransitive, and mixed verbs.



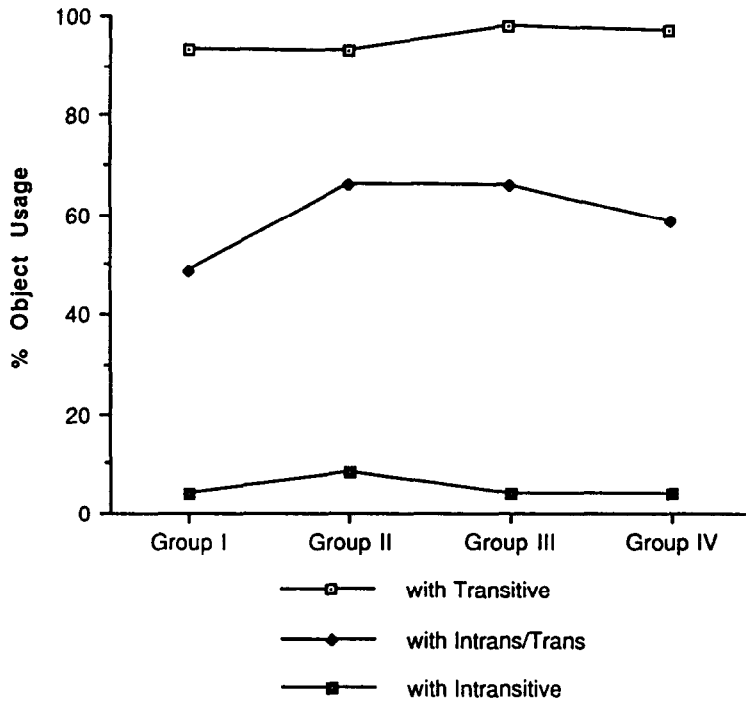
of mixed verbs decreases from .35 (s.d. = .16) in Group I to .29 (s.d. = .10) in Group II; there is another decrease from .33 (s.d. = .11) in Group III to .25 (s.d. = .06) in Group IV. Third, the proportion of pure intransitive verbs decreases gradually from .20 (s.d. = .10) in Group I, to .11 (s.d. = .05) in Group II, to .09 (s.d. = .05) in Group III, to .06 (s.d. = .06) in Group IV.

As can be seen from Figure 6, the children were reasonably constant in their use of objects for verbs which were either pure intransitives or pure transitives, but they slightly increased their use of objects for verbs which could be intransitive or transitive. The children seldom used objects with pure intransitive verbs: Group I used objects 4% of the time (s.d. = 4); Group II 8% (s.d. = 5); Group III 4% (s.d. = 6); Group IV 4% (s.d. = 8).

The children consistently used objects with pure transitive verbs, even in Group I, where the highest individual percentage of omission was 14%. With pure transitive verbs, Group I used objects 93% of the time (s.d. = 5); Group II 93% (s.d. = 8). Group III 98% (s.d. = 2); Group IV 97% (s.d. = 4).

The children's use of objects with mixed verbs increases somewhat from Group

Figure 6. American children's production of objects with different verb types.



I (.49, s.d. = .13) to Group II (.66, s.d. = .30), and remains roughly at that level (Group III averages .66, s.d. = .27; Group IV averages .59, s.d. = .08).

Discussion

Performance factors of some type do appear to play a role in children's use of subjects. First, we find support for the prediction that children will use shorter VPs with full lexical subjects than pronominal subjects, and shorter VPs with pronominal subjects than no subjects. Although only 6 out of the 10 children in Groups I and II showed the predicted pattern, that is more than three times as many as would be expected by chance, given that there are six possible orderings. Further, in Groups III and IV, where we would predict a weaker performance limitation, only 3 of the 11 children show the same pattern – a chance effect. For full satisfaction, however, Group I should have displayed the pattern more than Group II, and that did not occur.

Second, the other performance measures also showed effects. MLU and age combined predicted use of subjects well. The fact that age is related to subject use almost as strongly as MLU suggests that some language-independent factors

are involved in children's use of subjects. Increased subject use would, all other things being equal, contribute to increased MLU, so that relation, although predicted, is not especially strong confirmation of a performance account. But subject use obviously cannot contribute to age, and thus the independent contribution of age is noteworthy.

More interesting still is the correlation between subject use and verb use. Subject and verb use are linked across the entire range of American children. There is some evidence that subject and verb use are also linked for the Italian children, though within a narrower range of subject and verb use, but because of the small number of Italian children it is difficult to assess. Some effect in both language groups was predicted, since both should be subject to production limitations. When the American correlation between subject and verb use is put together with their much higher rate of subject use, it suggests that the American children know that verbs require subjects, but use them less at low MLUs because of performance limitations.

If American children do understand that their sentences require subjects, that raises the question of how they deal with direct objects. There is a marked asymmetry in how often Group I children use subjects and how often they use objects for pure transitive verbs. Since even Group I children produce objects when required over 90% of the time, we see little to support the hypothesis that they have *pro*, or any null or implicit argument, in object position. Bloom, Miller, and Hood (1975) similarly report a very high percentage of objects for verbs which appear to be pure transitives. Radford's (1990) report of null objects may be due to the lower age of the children he examined; alternatively, since he reports no quantitative data, the examples may be infrequent for each child.

Any theory, competence or performance, which predicts an asymmetry between subjects and objects between MLUs 1.5 and 2.0 is on firmer ground than one which predicts that they will be treated the same. The question then is whether the asymmetry means that Group I children fail to understand that subjects are required. When all the data are considered, the answer appears to be no.

The children's performance on verbs of different types, and their production of objects, is noteworthy in several respects. First, our data show that children typically do not use a verb unless they know how it subcategorizes with respect to objects. There are relatively few errors at any MLU of incorrectly using an object with a pure intransitive verb (and the errors that occur might be best described as omission of a preposition), and relatively few errors of incorrectly omitting an object with a pure transitive verb.

The most telling contrast is that between provision of objects with pure transitive verbs and mixed verbs. The children have those two classes separated: they provide objects much more frequently for pure transitive than mixed verbs, indicating that they recognize the difference between when an object is obligatory

and when optional. That suggests that the children are tracking parental input closely, since the only way of distinguishing a pure transitive verb like "hold" from a mixed verb like "eat" is by noticing whether adults consistently use an object after each of them.

Second, our data show that the use of optional objects in mixed verbs increases (from 49% to 66%) between Groups I and II. That increase seems best understood as a decrease in performance limitations. As the children can produce longer utterances, they provide more optional objects for mixed verbs. Third, the children increase their use of pure transitive verbs (from 45% to 69%) as development proceeds. That increase also seems best understood as a decrease in performance limitations. As the children can handle longer length, they increase their use of verbs which require objects.

We can now note that the child has different degrees of freedom with respect to subjects and objects. Subjects are always grammatically required (though they can be acceptably omitted in some circumstances) once the child's grammar exceeds the scope of a VP. Objects are always required for purely transitive verbs (both for grammaticality and acceptability), but the child has the option of using more intransitive and mixed verbs to get around the cognitive load that additional constituents would appear to impose (if there is a synonymous intransitive or mixed verb in the child's vocabulary).

Our data suggest that the Group I children make use of that option: they produce the lowest percentage of purely transitive verbs and the highest percentage of purely intransitive verbs; they also produce the lowest percentage of objects. With objects, the Group I children can grammatically avoid the burden objects impose. With subjects, the child does not have a comparable option. Our data on the increase in children's use of purely transitive verbs, which first occurs at the same time as the increase in children's subjects, suggests that children are able to increase the number of constituents that they can handle around MLU 2, and then again around MLU 4.

It may also be the case that semantic factors interact with cognitive load. The conceptual content of pure transitive verbs may be more complex than that of intransitive or mixed verbs. It is not clear, however, that there are any meaning differences which reliably correlate with transitivity, so this must remain speculation.

Summary

Study 5 has presented evidence that performance factors are important in children's production of subjects. Children in Groups I and II produced their longest VPs when no subject was present, next longest with a pronominal subject, and shortest with a full lexical subject. Subject use is highly correlated with MLU and age together, and with verb use. Children's production of purely transitive verbs

increases most between Groups I and II, as does their production of optional objects. The facts are best explained by performance theories.

General discussion

We have had three goals in this series of studies. The first was to determine the facts about use of subjects and related sentence constituents among young American children. The second was to decide, from those facts, when American children understand that their language requires subjects. The third was to evaluate competence and performance explanations of children's use of subjects.

At a factual level, our studies show that American children between MLU 1.50 and 2.00 (ages 1;10 to 2;2) use subjects in almost 70% of their utterances with verbs. Between MLU 2.00 and 3.00, subject use averages almost 90%, and around MLU 4.00 reaches 95%. Subject use is never lower than 84% after MLU 2. We first conclude that the children with MLU greater than 2 understand that subjects are required, and that neither the pro hypothesis nor the VP hypothesis holds for those children.

Our second conclusion is that American children understand even earlier, somewhere between MLU 1.5 and 2.0, that English requires subjects. A multi-factored performance explanation is a more tenable explanation of the children's behavior than a competence deficit. Group I, taken as a whole, provides no direct evidence in favor of the pro hypothesis (e.g., Hyams, 1986, 1987) or the VP hypothesis (e.g., Guilfoyle & Noonan, 1989; Kazman, 1988) and some evidence against. Features that, on the basis of either of those two competence explanations, would be expected to co-occur with inconsistent use of subjects, do not. Further, the children do not lack tense (O'Grady et al., 1989). If our subordinate clause data are taken at face value, there is support for Roeper and Weissenborn's (1990) suggestion that subordinate clauses are linked with obligatory subjects.

Let us review the evidence for the conclusion that, despite their inconsistent usage, Group I understands that sentences require overt subjects. Group I produced twice as many subjects as Italian children between the ages of 2;1 and 2;5 (Study 1), contrary to the pro hypothesis in Hyams (1986). They used pronouns for most of their subjects, again contrasting with Italian children, who used pronouns for a minority of their subjects (Study 1). They produced few expletive subjects, but as many as higher-MLU children did (Study 2). Not only did low-MLU American children differ from Italian children, but they also used subjects more often than the Japanese children studied by Mazuka et al. (1986). The subjects used by Group I look like real subjects rather than VP subjects, because the children consistently used nominatively case-marked pronouns in subject position (Study 2), contrary to the VP hypothesis.

Both competence hypotheses have predicted (or assumed) lack of modals, but the Group I children produced a few modals, and produced them more often than Italian children (Study 3). Further, there was no correlation between how frequently children produced subjects overall and how frequently they produced modals (Study 3). The only link between modals and subjects, one best accounted for by facts of English usage, was that children uniformly produced a subject if their utterance had a modal. Group I children also produced examples of infinitival *to*, past tense, third person singular present, and subordinate clauses (Study 4), all contrary to what the VP hypothesis would predict.

Some of the early infrequent uses of sentence elements that we have documented may be spurious, rather than genuine; both competence-deficit hypotheses would interpret them thus. The competence theories would then predict an abrupt quantitative or qualitative change in how children use those elements, a change which would occur at the same time as the large increase in use of subjects between Groups I and II. But the kind of patterned timing of emergence of elements that would be expected if either competence hypothesis were correct does not emerge from our detailed examination of the children's productions, with the possible exception of the difference between the lowest-MLU child and the remaining children. Almost all the observed changes in use appear to fall into four categories, none of which fit the predictions:

(1) The changes are gradual rather than abrupt. An example is the increase in number of types and tokens of modals.

(2) The changes occur at the wrong time. Examples are the increase in infinitival *to* between Groups II and III; over-regularization of the past tense between Groups II and III; the increase in third person singular *s* between Groups III and IV; and the increase in subordinate tensed clauses between Groups II and III.

(3) The changes are predicted by performance accounts rather than by competence accounts. Examples are the higher than chance usage of longer VPs with no subject than with a pronominal or lexical subject; the correlation between MLU, age, and subject use; the correlation between verb use and subject use; the increase in purely transitive verbs; the increase in objects provided for mixed verbs.

(4) The changes are independent of any theory at present (which is natural, as there will be some simultaneous changes that are due to other grammatical developments). Examples of independent changes occurring between Groups I and II are the increase in use of pronominal subjects (already high in Group I), also evident in the Italian children; the increase in the use of semi-Aux's between Groups I and II.

While we have concluded that performance explanations account better for the overall pattern of the children's development, we note that the lowest-MLU child, at MLU 1.53, provided some evidence for the VP hypothesis. This child not only

used subjects less than any other child, but had few personal pronouns as subjects, no modals, no semi-Aux's, and no regular past tense. There were, however, two infinitives, one third person singular, and one subordinate clause. While not presenting a completely consistent picture, this child came closer to fitting the VP description than did any other. Accordingly, the initial stage for all children may be a version of the VP grammar illustrated in Figure 1 (e.g., Guilfoyle & Noonan, 1989).

Our suggestion, however, is that the child's first syntax is a skeletal mobile consisting of the entire tree in Figure 1 (minus modals), with the nodes unfilled and with the order of SPEC, head, and complement left free. We favor this version on the assumption that all the elements in the skeletal tree (except modals) are universal and innate. The developing child learns how to lexicalize the nodes, learns what the proper phrase orders are in her particular language, and learns what empty categories exist in her language. The child in our study at MLU 1.53 is clearly deficient in knowing how to lexicalize different sentence elements; whether the child also lacks the full skeletal structure we cannot say.

With respect to *pro*, the developing child will wait for evidence before projecting that empty category, and will require a licenser of some sort for it. The American child does receive some misleading input, namely utterances without subjects. But there is no evidence that American children are ever seriously misled by such utterances to project *pro*. Further, it seems unlikely that the American child can start off very much in error if she already has the correct notion of subjects before MLU 2.0.

The Italian child receives a great deal of information that empty subjects are possible, in the form of utterances without subjects (on the basis of Bates's (1976) data, about 50% of adult input) and in the form of "perfect" verb endings. Unlike the American child, the Italian child not only receives evidence that null subjects are possible, but evidence about how they are licensed. Since the empty subject's identity can be read off the verb endings, identification is guaranteed, and if identification is guaranteed, so is licensing. (Whether such reasoning will actually guarantee the correct outcome depends in part on whether it is universally true that languages with completely transparent verb endings have *pro* subjects.) In the case of the child learning Japanese, Chinese, and many other "null subject" languages, we can say nothing further at present, because the nature of the empty subject is not clear, nor how the grammar licenses the empty category.

All other things being equal, we want to construct an acquisition mechanism which projects the minimum number of incorrect hypotheses (Guilfoyle & Noonan, 1989; Lebeaux, 1989), because grammar reorganization is costly. At the same time, we have to account for the errors the child produces. An incomplete grammar, representing absence of knowledge, rather than an incorrect grammar, representing false information, is one way to achieve both. That is the attraction of the VP hypothesis: it postulates a limited grammar which can serve as a correct

foundation for later grammatical development, rather than an incorrect grammar which must be revised.

Recall that the pro hypothesis had no theory-internal basis for starting the child off with the null subject value of the null subject parameter (Hyams, 1987): it was in response to presumed facts about children's early productions – presumptions which our data show were incorrect. While the subset principle has sometimes been invoked as a reason for starting the child off with the *non-null* value, that reasoning receives no empirical support and also appears logically flawed (Valian, 1989, 1990a).

Our suggestion – a skeletal grammar – is similar to the VP hypothesis in proposing that the child's initial knowledge is incomplete rather than in direct commissive error. It differs from the VP hypothesis in invoking no mechanism other than learning as the effector of changes. The child learns, via the evidence to which she is exposed, how tense and agreement are coded in her language, how elements are ordered within a phrase, what empty categories exist, and so on. As the child learns, her performance mechanism is concurrently developing, affording longer utterances and more complete expression of what the child knows. Performance factors must play a role in production, regardless of the child's target language. The question is how those factors interact with children's knowledge to produce the observed patterns of use. We suggest that the inconsistent use of subjects for most of the children below MLU 2.0 is due to performance limitations, and have presented evidence to that effect. The work for the future is to develop performance models that will allow us to isolate both what children know, and what they can express.

References

- Antinucci, F., & Parisi, D. (1973). Early language acquisition: A model and some data. In C.A. Ferguson & D.I. Slobin (Eds.), *Studies of child language development*. New York: Holt, Rinehart, and Winston.
- Bates, E. (1976). *Language and context*. New York: Academic Press.
- Bellugi, U. (1967). *The acquisition of negation*. Unpublished Ph.D. dissertation, Harvard University.
- Blake, J., Quartaro, G., Austin, W., & Vingilis, E. (1989). *Does memory span constrain language complexity?* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Kansas City, MO.
- Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Bloom, L. (1991). *Language development from two to three*. Cambridge: Cambridge University Press.
- Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development*, 40.
- Bloom, L., Miller, P., & Hood, L. (1975). Variation and reduction as aspects of competence in language development. In A. Pick (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 9, pp. 3–55). Minneapolis: University of Minnesota Press.

- Bloom, L., Tackeff, J., & Lahey, M. (1984). Learning *to* in complement constructions. *Journal of Child Language*, 11, 391–406.
- Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, 21, 491–504.
- Borer, H. (1989). Anaphoric AGR. In O. Jaeggli & K.J. Safir (Eds.), *The null subject parameter*. Dordrecht: Kluwer.
- Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 123–172). Dordrecht: Reidel.
- Bowerman, M. (1973). *Early syntactic development: A cross-linguistic study with special reference to Finnish*. Cambridge: Cambridge University Press.
- Bowerman, M. (1978). Semantic and syntactic development. In R.L. Schiefelbusch (Ed.), *Bases of language intervention* (pp. 97–189). Baltimore: University Park Press.
- Braine, M.D.S. (1974). Length constraints, reduction rules, and holophrastic processes in children's word combinations. *Journal of Verbal Learning and Verbal Behavior*, 13, 448–456.
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Brown, R., Cazden, C., & Bellugi, U. (1973). The child's grammar from I to III. In C.A. Ferguson & D.I. Slobin (Eds.), *Studies of child language development* (pp. 295–333). New York: Holt, Rinehart and Winston.
- Brown, R., & Fraser, C. (1963). The acquisition of syntax. In C.N. Cofer & B.S. Musgrave (Eds.), *Verbal behavior and learning* (pp. 158–197). New York: McGraw-Hill.
- Budwig, N. (1985). "I, me, my and 'name'": children's early systematizations of forms, meanings and functions in talk about the self. *Papers and Reports on Child Language Development*, 24, 30–37.
- Budwig, N. (1989). The linguistic marking of agentivity and control in child language. *Journal of Child Language*, 16, 263–284.
- Chi, M.T. (1978). Knowledge structures and memory development. In R.S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.
- de Villiers, J.G., & de Villiers, P.A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, 2, 267–278.
- Ervin, S.M. (1964). Imitation and structural change in children's language. In E.H. Lenneberg (Ed.), *New directions in the study of language*. Cambridge, MA: MIT Press.
- Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, 30.
- Greenfield, P.M., & Smith, J.H. (1976). *The structure of communication in early language development*. New York: Academic Press.
- Guilfoyle, E. (1984). *The acquisition of tense and the emergence of lexical subjects in child grammars of English*. Boston University Conference on Language Development, Boston.
- Guilfoyle, E., & Noonan, M. (1989). *Functional categories and language acquisition*. Unpublished manuscript, McGill University.
- Huang, C.T.J. (1989). pro-drop in Chinese: A generalized control theory. In O. Jaeggli & K.J. Safir (Eds.), *The null subject parameter*. Dordrecht: Kluwer.
- Hyams, N.M. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.
- Hyams, N.M. (1987). *The setting of the null subject parameter: A reanalysis*. Boston University Conference on Language Development, Boston.
- Jaeggli, O., & Safir, K.J. (1989). The null subject parameter and parametric theory. In O. Jaeggli & K.J. Safir (Eds.), *The null subject parameter*. Dordrecht: Kluwer.
- Kazman, R. (1988). *Null arguments and the acquisition of Case and Infl*. Boston University Conference on Language Development, Boston.
- Kuczaj, S.A. (1981). More on children's initial failure to relate specific acquisitions. *Journal of Child Language*, 8, 485–487.
- Lebeaux, D. (1987). Comment on Hyams. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 23–39). Dordrecht: Reidel.

- Lebeaux, D. (1989). *A formal syntax of telegraphic speech*. Boston University Conference on Language Development, Boston.
- Limber, J. (1976). Unravelling competence, performance and pragmatics in the speech of young children. *Journal of Child Language*, 3, 309–318.
- Lust, B., & Chien, Y.-C. (1984). The structure of coordination in first language acquisition of Mandarin Chinese: Evidence for a universal. *Cognition*, 17, 49–83.
- Marcus, G.F., Ullman, M., Pinker, S., Hollander, M., Rosen, T.J., & Xu, F. (1990). *Overregularization*. Occasional Paper no. 41, Center for Cognitive Science, MIT, Cambridge, MA.
- Mazuka, R., Lust, B., Wakayama, T., & Snyder, W. (1986). Distinguishing effects of parameters in early syntax acquisition: A cross-linguistic study of Japanese and English. *Papers and Reports on Child Language Development*, 25.
- Meisel, J.M. (1990). *INFL-lection: Subjects and subject-verb agreement in early child language. Evidence from simultaneous acquisition of two first languages: German and French*. Unpublished manuscript, University of Hamburg.
- O'Grady, W., Peters, A.M., & Masterson, D. (1989). The transition from optional to required subjects. *Journal of Child Language*, 16, 513–529.
- Olson, G.M. (1973). Developmental changes in memory and the acquisition of language. In T.E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 145–157). New York: Academic Press.
- Peters, A.M. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry*, 20, 365–424.
- Radford, A. (1988). *Transformational grammar: A first course*. Cambridge: Cambridge University Press.
- Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Oxford: Blackwell.
- Riemsdijk, H. van, & Williams, E. (1986). *Theory of grammar*. Cambridge, MA: MIT Press.
- Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry*, 17, 501–557.
- Roeper, T., & Weissenborn, J. (1990). How to make parameters work: Comments on Valian. In L. Frazier & J. de Villiers (Eds.), *Language processing and language acquisition* (pp. 147–162). Dordrecht: Kluwer.
- Stromswold, K. (1989). *Using naturalistic data: Theoretical and methodological issues*. Boston University Conference on Language Development, Boston.
- Ullman, M., Pinker, S., Hollander, M., Prince, A., & Rosen, T.J. (1989). *Growth of regular and irregular vocabulary and on the onset of overregularization*. Boston University Conference on Language Development, Boston.
- Valian, V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology*, 22, 562–579.
- Valian, V. (1989). *Positive evidence, indirect negative evidence, parameter-setting, and language learning*. Unpublished manuscript, Hunter College, New York.
- Valian, V. (1990a). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition*, 35, 105–122.
- Valian, V. (1990b). Logical and psychological constraints on the acquisition of syntax. In L. Frazier & J. de Villiers (Eds.), *Language processing and language acquisition* (pp. 119–145). Dordrecht: Kluwer.
- Valian, V. (in press). *Parental replies: Linguistic status and didactic role*. MIT Press/Bradford, Cambridge, MA.