Developmental Psychology
1995, Vol. 31, No. 2, 180–197

0012-1649/95/$3.00

Negative Evidence on Negative Evidence

James L. Morgan and Katherine M. Bonamo
Brown University

Lisa L. Travis
University of Minnesota

Previous work has shown that recasts may be contingent responses to children's early ungrammaticality. On this basis, it has been claimed that recasts provide negative evidence, thereby offsetting the need for linguistic constraints in theories of acquisition. This study explores whether children exploit negative evidence putatively provided by recasts by examining whether parental recasts are associated with children's recovery from particular overgeneralization errors. Data from longitudinal investigations of 2 common syntactic errors reveal that recasts are related to children's subsequent grammaticality. However, contrary to what would be expected if recasts serve as corrections, the data show that recasts are negative leading indicators of grammaticality. Finally, *correction* and *negative evidence* are examined and are shown to be nonequivalent. Therefore, corrections in whatever form they might exist can offset only a limited subset of proposed innate constraints on language acquisition.

In the course of development, children may at times adopt grammars that appear to be overgeneral with respect to the language they are acquiring. For example, children learning English may alternate between using and omitting articles preceding singular common count nouns like *cat*; mature English allows only a subset of these possibilities, requiring articles to be used in this context. Accounting for how children avoid or recover from such overgeneralizations is a central conundrum for theories of language acquisition. Positive evidence that children receive—examples of grammatical sentences—provides no basis for restricting overgeneralizations. If children receive only positive evidence, the solution to this conundrum must be sought by appealing to structured properties of the mind. A logical alternative is that children receive negative evidence: corrective feedback providing information that certain sentences are not acceptable. Such evidence would allow children to recover from overgeneralizations through simple learning mechanisms.

From the inception of the contemporary debate on nature and nurture in children's language acquisition, the issue of negative evidence has been a focus of contention. Proponents of theories emphasizing environmental control over acquisition have asserted that differential responses to children's grammatical and ungrammatical utterances are crucial for fostering learning of syntax:

> In teaching the young child to talk, the formal specifications upon which reinforcement is contingent are at first greatly relaxed. Any response which vaguely resembles the standard behavior of the community is reinforced. When these begin to appear frequently, a closer approximation is insisted upon. In this manner very complex verbal forms may be reached. (Skinner, 1957, pp. 29–30)

In contrast, proponents of theories emphasizing control of properties of the mind over acquisition have expressed skepticism concerning the availability of such differential responses in children's environments:

> I have been able to find no support whatsoever for the doctrine of Skinner and others that slow and careful shaping of verbal behavior through differential reinforcement is an absolute necessity. If reinforcement theory really requires the assumption that there be such meticulous care, it seems best to regard this simply as a reductio ad absurdum argument against this approach. (Chomsky, 1959, pp. 42–43)

In theories inclining toward more nativist views, the *no negative evidence assumption* is cited prominently as a fundamental premise (Baker, 1979; Lightfoot, 1991; Pinker, 1984, 1989; Wexler & Culicover, 1980). Such theories have proceeded to marshal arguments (including observations that children commit relatively few errors and acquire a highly abstract grammatical system on the basis of limited and perhaps impoverished exposure to language) for detailed and explicit proposals concerning innate properties of mind underlying language acquisition. The empirical basis of the no-negative-evidence assumption on which much of this theoretical edifice is founded, however, is rather slim: a scattering of anecdotes (Braine, 1971; McNeill, 1966), a study by Brown and Hanlon (1970) showing that explicit parental feedback (such as saying something is "Right" or "Wrong") is not associated with children's grammaticality, and, most compellingly, a traditional absence of clear demonstrations of how negative evidence might be manifested in children's language input.

James L. Morgan and Katherine M. Bonamo, Department of Cognitive and Linguistic Sciences, Brown University; Lisa L. Travis, Institute of Child Development, University of Minnesota.

This research was supported in part by Biomedical Research Support Grant 5-27350 to Brown University and by National Institute of Child Health and Human Development Grant T32-HD07151 to the Center for Research in Learning, Perception, and Cognition, University of Minnesota. We thank Christopher Bingham, Pedro Gozalo, Ulf Grenander, and Bernie Saffran for statistical consultation and advice.

Correspondence concerning this article should be addressed to James L. Morgan, Department of Cognitive and Linguistic Sciences, Brown University, Box 1978, Providence, Rhode Island 02912. Electronic mail may be sent to JAMES_MORGAN@BROWN.EDU.

Beginning with Hirsh-Pasek, Trieman, and Schneiderman (1984), several recent articles have reported that different types of adult responses may in fact be associated with children's grammatical and ungrammatical utterances. Bohannon and Stanowicz (1988), Demetras, Post, and Snow (1986), Hirsh-Pasek et al. (1984), Morgan and Travis (1989), and Penner (1987) were in concordance that in early stages of language acquisition children's ungrammatical utterances (e.g., "See ball") are more likely to elicit expansions or recasts (e.g., "See the ball") than are their grammatical utterances.[1] Children's grammatical utterances, on the other hand, are more likely to elicit exact imitations or conversational move-ons. These results support the possibility that recasts—reformulations of children's utterances—may manifest negative evidence, serving as corrective feedback. This has sparked further theoretical interest because, in Bohannon and Stanowicz's (1988, p. 684) words "if children's conversational partners provide some form of corrective feedback (i.e., negative evidence), then many of the innate linguistic constraints recently proposed would become unnecessary."

This last conclusion, as has been argued by Gordon (1990), Pinker (1989), and others, is premature: Evidence that adult responses are related in different ways to grammatical and ungrammatical child utterances is not in itself sufficient to settle the negative evidence argument. Parents could correct children's errors religiously, but it would matter little if children fail to apply those corrections in constructing their grammars. Child-internal processes of *discrimination, interpretation,* and *reevaluation* must be involved in utilizing corrective feedback, and none of these is trivially accomplished.[2]

First, if recasts are to serve as corrections in acquisition, children must be able to discriminate recasts from other types of adult responses. Note, however, that there is nothing distinctive about the forms of recasts. Whether an adult utterance qualifies as a recast depends solely on its relation to a preceding child utterance. This stands in contrast to the explicit feedback examined by Brown and Hanlon (1970), which was manifested by a fixed set of forms. To discriminate recasts, children must therefore be able to compare their utterances to the responses these evoke and apply some metric to the comparison to determine whether the right sort of near-match occurred. Ideally, the metric would include parental utterances that added to, subtracted from, or rearranged words in child utterances for grammatical purposes while excluding parental utterances that introduce such changes for other purposes (e.g., providing a more complete description of a scene or altering the focus of a sentence). Of course, children are not in a position to formulate an ideal metric, for this would entail a priori grammatical knowledge of the language being learned. How children might formulate an alternative metric is unknown. Regardless of the metric chosen, however, the resulting category of recasts will include some proportion of adult responses that do not embody grammatical corrections (or exclude some proportion of adult responses that do embody grammatical corrections, or both), failing to capture all of the possible information in input.

Second, children must be able to interpret recasts as manifesting corrective feedback. Again unlike the explicit feedback studied by Brown and Hanlon (1970) and Hirsh-Pasek et al. (1984), nothing intrinsic to the forms or meanings of recasts is corrective. Moreover, evidence available from input may indicate that recasts are *not* corrections. To see why this is so, consider the following hypothetical exchanges.

1. Child: What that can do?
   Parent: What can that do?
2. Child: He turned the light on.
   Parent: Yes, he turned on the light.

In both cases, the parental response recasts the child's utterance. Suppose the child has hypothesized that recasts are corrections. One might suppose that in instances like Example 2 this hypothesis would have little cost: Additional positive examples would show that sentences like "He turned the light on" are in fact permissible. But note that such additional positive examples also provide direct evidence that recasts are not corrections. Parents also recast their own utterances (Snow, 1972), providing more evidence of the same sort. To maintain the recasts-as-corrections hypothesis in the face of evidence to the contrary, children would seem to need just the sort of powerful biases that environmentalist explanations seek to avoid.

Third, children must be able to reevaluate their current grammars in light of the corrective feedback supplied by recasts, devising modifications that more closely approximate the target language. Here, recasts would seem to be advantaged in relation to explicit feedback because they illustrate to children the manners in which their utterances were defective. But recasts do not inform children how to formulate appropriate grammatical modifications any more than do other forms of positive evidence: This is a matter left to be solved by induction.

We consider the issue of reevaluation to be of paramount importance in resolving the negative evidence argument. If it can be shown that recasts catalyze changes in children's grammars—specifically, driving children to discard ungrammatical forms of constructions—then it may be presumed that children succeed in solving the problems of discrimination and interpretation, the difficulties cited above notwithstanding. On the other hand, if recasts do not impel children to discard ungrammatical constructions, they cannot be corrections, and the issues of discrimination and interpretation become moot.

Several previous studies have considered the role of recasts in acquisition; overall, the results of these studies are quite equivocal. Studies correlating global incidence of recasts or expansions with general measures of language growth (Gleitman, Newport, & Gleitman, 1984; Hoff-Ginsberg, 1985; Scarborough & Wycoff, 1986) have typically found weak or nonexistent associations. In contrast, studies examining short-term impact of recasts of specific constructions on subsequent use of those constructions have found positive effects, whether the recasts are supplied as an experimental manipulation (Baker & Nelson, 1984; Nelson, 1977; Nelson, Carskaddon, & Bonvillian, 1973) or are observed in natural parent–child interactions (Farrar, 1990).

---

[1] As Valian (in press) pointed out, however, the exact degree of concordance is impossible to ascertain because these articles have used coding categories with varying degrees of inclusiveness for both adult and child utterances.

[2] Pinker (1989) provided an alternative analysis of the conditions that must be fulfilled if recasts are to provide correction.

The available evidence leaves undecided the issue of whether recasts provide negative evidence and serve as corrections. On one hand, the negative results may simply reflect the choice of inappropriate time lags or irrelevant (for present purposes) measures of growth. Supposing that recasts were corrections, there is no particular reason to expect contingent growth in, say, mean length of utterance (MLU). Growth in MLU indexes the relative preponderance of complex constructions and the relaxation of length constraints in production, neither of which need be directly related to the disappearance of specific ungrammatical forms. On the other hand, the positive results may only reflect children's incorporation of new alternatives in their grammars. Even if recasts lead to absolute increases in use of grammatical forms of particular constructions, this does not imply that they lead to decreases in use of ungrammatical forms of those same constructions. For ascertaining the corrective potential of recasts, measures of proportional increases in use of specific grammatical forms would be more relevant.

In the studies reported here, we investigated whether recasts are associated with the later suppression of the ungrammatical forms that are recast. We examined longitudinal effects of recasts on children's acquisition of two specific constructions: article use with singular count nouns and auxiliary use in *Wh* (who, what, when, where, how, and why) questions. In the first instance, the error of interest involves omission of the obligatory article; in the second, the error of interest involves omission or misplacement of the auxiliary verb following a fronted Wh word. We chose these constructions for several reasons. First, neither of these constructions, as it is manifested in English, is universal; effects of input may be more likely to appear with respect to language-specific constructions (cf. Newport, Gleitman, & Gleitman, 1977). Second, these are both common constructions, and children's errors are also quite common, thus ensuring frequent opportunities to observe recasts. Third, in both instances, children alternate in using the relevant ungrammatical and grammatical forms, so that these errors are examples of overgeneralizations. Finally, these two errors involve different grammatical processes (morphology vs. movement), and these constructions are acquired in different epochs of development. These differences afford a degree of generalization.

The general hypothesis we wished to test is straightforward. If recasts provide negative evidence and serve as corrections, we expect to see some positive relation between the incidence of recasts and the disappearance of syntactic errors. Unfortunately, no explicit theory of how corrections are used in acquisition has been forthcoming, and the specific form of expected relation is unclear. We therefore conducted exploratory analyses on two time scales. The first of these concerned immediate effects of individual recasts on children's grammaticality and self-repairs, whereas the second concerned mass effects of recasts on proportional grammaticality over lags ranging from 2 to 12 weeks.

Study 1

In English, singular count nouns, such as *dog* or *table,* must be accompanied by one of four types of specifiers: demonstratives (*this* or *that*), singular or generic quantifiers (*one, another, each, no, any,* or *every*), possessives (*my, your, John's,* and so forth), or articles (*a, an,* or *the*). Articles occur most often, being among the most frequent of English words (Kucera & Francis, 1967). Nevertheless, articles are conspicuously absent in the earliest stages of acquisition; their appearance is one of the hallmarks of Stage 2 of acquisition (Brown, 1973). Articles were among the 14 morphemes whose acquisition was studied by Brown (1973) and deVilliers and deVilliers (1973). Maratsos (1976) investigated children's acquisition of semantic distinctions between definite and indefinite articles. Here, we were concerned only with children's observation of the syntactically obligatory status of articles with singular count nouns, regardless of whether children's article choices were semantically appropriate or not. Our fundamental question was whether parental recasts of children's article omissions are associated with lower future levels of such errors.

Method

The Adam, Eve, and Sarah transcripts, which have been described by Brown (1973) and several additional papers referenced there, provided our database. These transcripts were made available to us under the auspices of the Child Language Data Exchange System (MacWhinney & Snow, 1985).

Adam's, Eve's, and Sarah's use of articles was tracked from the point when article use began to the point when articles consistently appeared in at least 90% of obligatory contexts (the criterion for acquisition originally used in Brown, 1973). A computer search established the initial point, and information from Brown (1973) was used to establish the final point. Included in our tabulations were transcripts from Adam between the ages of 2 years 6 months to 3 years 6 months (Transcripts 7 to 35), Eve from 1 year 6 months to 2 years 4 months (Transcripts 1 to 20), and Sarah from 2 years 3 months to 3 years 8 months (Transcripts 1 to 71).

Concordances were prepared by computer for each child over the appropriate range of transcripts. All words that could be singular count nouns were culled from these lists. Then, two exhaustive files were prepared. The first, "ungrammatical" file was built by a computer search that identified all instances of possible singular count nouns that were not immediately preceded by articles. The second, "grammatical" file was built by a computer search that identified all instances of possible singular count nouns preceded by an article in the same utterance (with arbitrary numbers of intervening words allowed). Entries in each file consisted of the entire conversational turn containing the target noun use or uses (with the possible nouns highlighted) and the following adult conversational turn, with notation of the location in the transcript of each turn.

Each file was coded in two passes. In the first pass, possible noun uses were displayed one by one and the coder decided whether each should be included in later analyses. Examples were excluded if the target word was not a singular count noun in the child's utterance (e.g., if *watermelon* were used as a mass noun rather than a count noun, or if *record* were used as a verb rather than a noun). Some ungrammatical examples were misclassified (e.g., because one or more modifiers intervened between an article and a noun, or because the noun was preceded by some nonarticle specifier, such as *another*); these were excluded. Potential grammatical examples were excluded if the preceding article identified by the computer actually modified another noun. Assignment of examples to the grammatical category was not based on whether the correct article had been used but merely whether any article had been used.

In the second pass, the adult response to each grammatical or ungrammatical singular count noun use was assigned to one of 10 mutually exclusive categories. A *no response* was coded if one or more child

Table 1
Types of Responses to Children's Singular Count Noun Uses

Category	Child utterance	Adult response
Minimal recast	Sun fall down.	The sun fall down? Well, yes, it looks like it.
Expanded recast	Lady hat on.	The lady has a hat on?
Related recast	I want piece of it.	Here it is. Here's your piece.
Other recast	Dis a tape recorder.	That's a tape recorder?
Nonrecast expansion	Take a bath!	I don't want to take a bath!
Imitation	I don't got nose bleed.	I don't got nose bleed.
Clarification question	It's a motorboat duck?	What? Say it again.
Confirmation question	I eating a pear.	Are you saying pear?
Move-on	Mama isn't boy, he a girl.	That's right.

utterances immediately followed the utterance containing the target and if no utterance in the next adult conversational turn was explicitly related (through complete or partial imitation) to the noun-containing utterance. An initial code of *expansion* was given if any utterance in the adult conversational turn imitated one or more phrases of the noun-containing utterance and added or rearranged one or more words or morphemes. Several subcategories of expansions were coded. If an adult utterance expanded an ungrammatical noun-containing utterance by adding only an article, it was coded as a *minimal recast*.[3] If an adult utterance expanded an ungrammatical noun-containing utterance by adding an article and additional material, it was coded as an *expanded recast*. If an adult utterance expanded an ungrammatical noun-containing utterance by adding a nonarticle specifier, it was coded as a *related recast*. If an adult utterance expanded a grammatical or ungrammatical noun-containing utterance and corrected some nonarticle error, it was coded as an *other recast*. If an adult utterance expanded a grammatical or ungrammatical noun-containing utterance without correcting any syntactic error, it was coded as a *nonrecast expansion*. An *imitation* was coded if the adult exactly repeated the noun-containing utterance (reversals in pronouns and diectic terms notwithstanding). A *clarification question* was coded if any adult utterance had the force of requesting the child to repeat part or all of the noun-containing utterance. A *confirmation question* was coded if the adult reply to the noun-containing utterance was a yes–no question pertaining to the linguistic content of the utterance. Finally, a *move-on* was coded if none of the other categories was applicable. Only one response type was assigned to each target; in cases of multiple possible codes for a given response, the categories were given priority in the order listed here. Table 1 provides examples of each response type.

J.L.M. and K.B. shared responsibility for coding the adult responses, each coding about half of the transcripts. Approximately one quarter of each coder's half was blind coded by the other coder for reliability. Agreement was greater than 85% for all transcripts.

Results and Discussion

Our data set included more than 9,000 singular count noun uses. We found slightly more grammatical article-plus-singular-count-noun uses than ungrammatical bare-singular-count-noun uses for each child. Adam's data included 2,253 grammatical and 2,077 ungrammatical noun uses, Eve's data included 1,263 grammatical and 1,215 ungrammatical noun uses, and Sarah's data included 1,360 grammatical and 1,167 ungrammatical noun uses. The proportion of grammatical noun use in each transcript is shown for each of the 3 children in Figure 1.[4]

We were unable to discern any principled basis for the early restrictions in article use. Articles were sometimes used and sometimes not used for nouns in both subject and object position. Nor were articles used with nouns in certain semantic fields but not others. To see if article use freely alternated with nonuse, we searched for instances in which the children had used the same noun both with and without an article in the same conversational turn. In Adam's transcripts, we found 108 examples, including 59 instances of change from article use to nonuse and 49 instances of change in the opposite direction. In Eve's transcripts, we found 47 examples, including 23 instances of change from article use to nonuse and 24 instances of change in the opposite direction. In Sarah's transcripts, we found 25 examples, including 15 instances of change from article use to nonuse and 10 instances of change in the opposite direction. None of these differences is significant by the sign test. Thus, children appear to treat articles as optional, rather than obligatory, at least in the earlier periods of acquisition, suggesting that article omission is truly an error of overgeneralization.

Frequencies of responses to children's grammatical and ungrammatical noun uses are shown in Figure 2. Because the number of confirmation questions was so low, these were combined with clarification questions. Chi-square analyses of these data (combining all types of recasts and nonrecast expansions in the superordinate category *expansion*) showed significant associations between response type distribution and grammaticality for all 3 children: Adam, $\chi^2(4, N = 4{,}330) = 158.2$, $p < .01$; Eve, $\chi^2(4, N = 2{,}478) = 50.4$, $p < .01$; Sarah, $\chi^2(4, N = 2{,}527) = 147.1$, $p < .01$. For all 3 children, expansions were more likely to occur in response to ungrammatical noun uses, whereas move-ons were more likely to occur in response to grammatical noun uses. This pattern of results replicates that found in all previous studies of differential adult responses to grammatical and ungrammatical child utterances.

In one important respect, however, our data diverge from those reported in previous investigations. Article omission errors appear to be recast at a higher rate than are syntactic errors generally. Bohannon and Stanowicz (1988), who used a coding category similar to ours, found that 15% of their subjects' ill-

[3] Inclusion of subcategories of expansions in our coding system begs the question of whether children are capable of discriminating among these types of responses. In particular, recasts as defined here include only syntactic corrections; classifying responses as recasts therefore entails preexisting knowledge of grammar, which children lack. Nevertheless, our recast categories comprise those responses that most transparently encode corrective feedback. We included these categories in our coding system to provide the most generous possible circumstances for testing whether potential negative evidence in input affects acquisition.

[4] Each transcript for Adam and Eve included 2 hr of conversation gathered within a 2-week interval. Each of Sarah's original transcripts included 0.5 hr of conversation; the interval between her transcripts was 1 week. Here, we combined each successive pair of Sarah's transcripts into a single session to increase both the stability of her data and the comparability of the data across the 3 children.

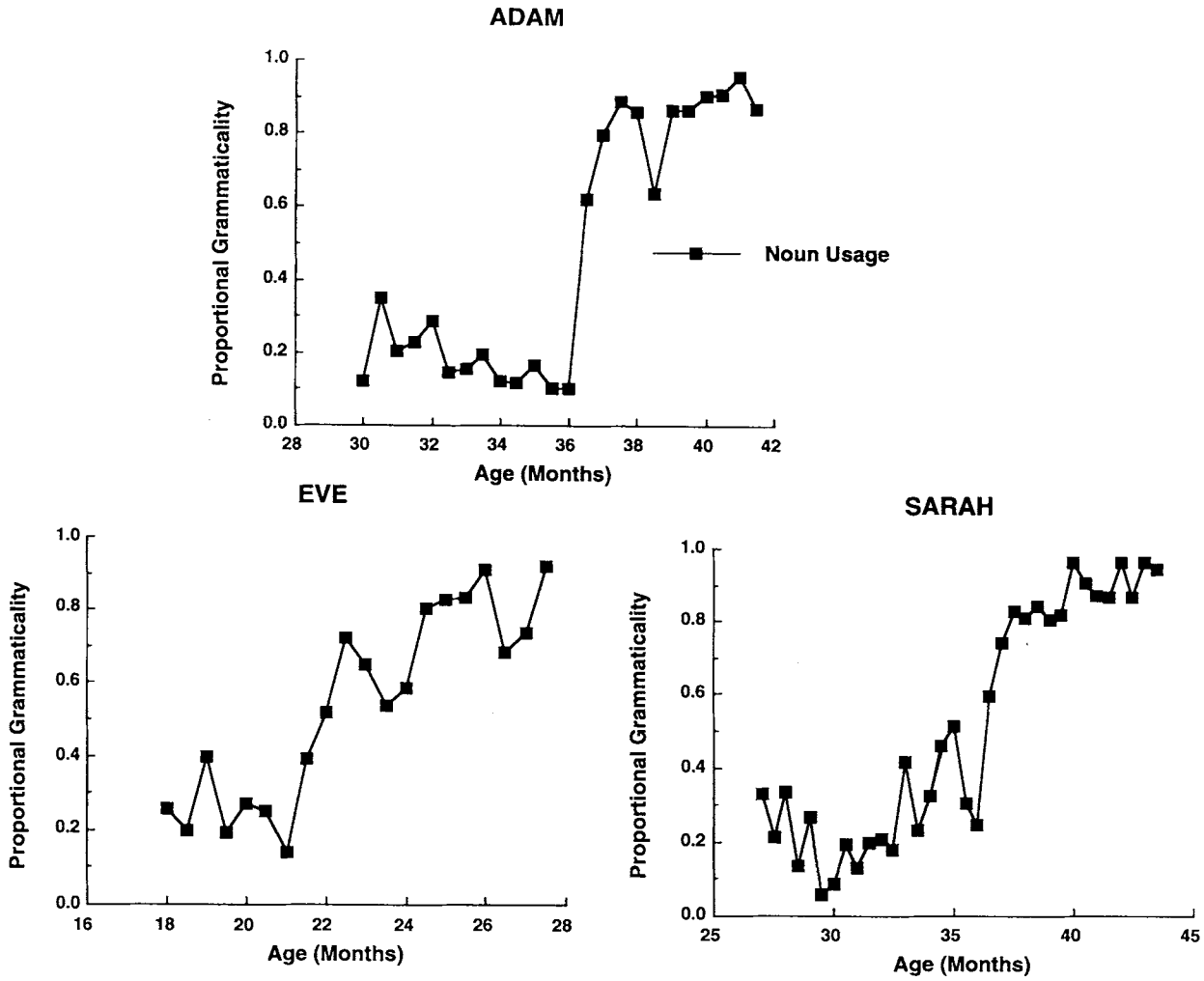

Figure 1. Proportions of children's grammatical article use.

formed utterances were recast. Other studies using more inclusive categories have found somewhat higher rates; for example, Hirsh-Pasek et al. (1984) found that 21% of their youngest subjects' ill-formed utterances were repeated (their repetition category included recasts, expansions, and exact imitations). In contrast, we found that specifier recasts (i.e., minimal recasts, expanded recasts, and related recasts) formed an especially high proportion of actual adult responses to ungrammatical noun uses: 34.8% for Adam, 41.4% for Eve, and 35.3% for Sarah. Farrar (1992), approaching the problem from a different perspective, provides some corroborating evidence. He reports that over 44% of the recasts he observed included recast articles, a far higher proportion than that observed in connection with any of the other morphemes he studied.

Article omission errors occur with high frequency and constitute a substantial fraction of all syntactic errors. Thus, one implication of the finding that they are also recast at particularly high rates is that at least some other types of syntactic errors must be recast quite infrequently. This raises questions about both the uniformity and robustness of recasts in input. A second implication is that effects of recasts ought to be especially apparent in children's use of articles. Conversely, failure to find effects of recasts on article use would be particularly difficult to reconcile with the hypothesis that corrections provided by recasts plays a critical role in acquisition.

There are, of course, several ways in which effects of recasts could be manifest. Recasts could incrementally reinforce use of particular grammatical forms, resulting in their immediate increased use. Recasts could make children more self-conscious about their grammar, prompting them to make more self-corrections. Recasts could competitively strengthen grammatical forms, thereby gradually driving out ungrammatical forms and increasing overall grammaticality. Below, we describe how we tested each of these possibilities, beginning with analyses examining relations among successive utterances and proceeding to more macroscopic analyses examining influences on overall grammaticality across time.

Short-term effects of recasts. Farrar (1992) has recently reported that 23-month-olds are two to three times more likely to imitate grammatical morphemes in corrective recasts than the

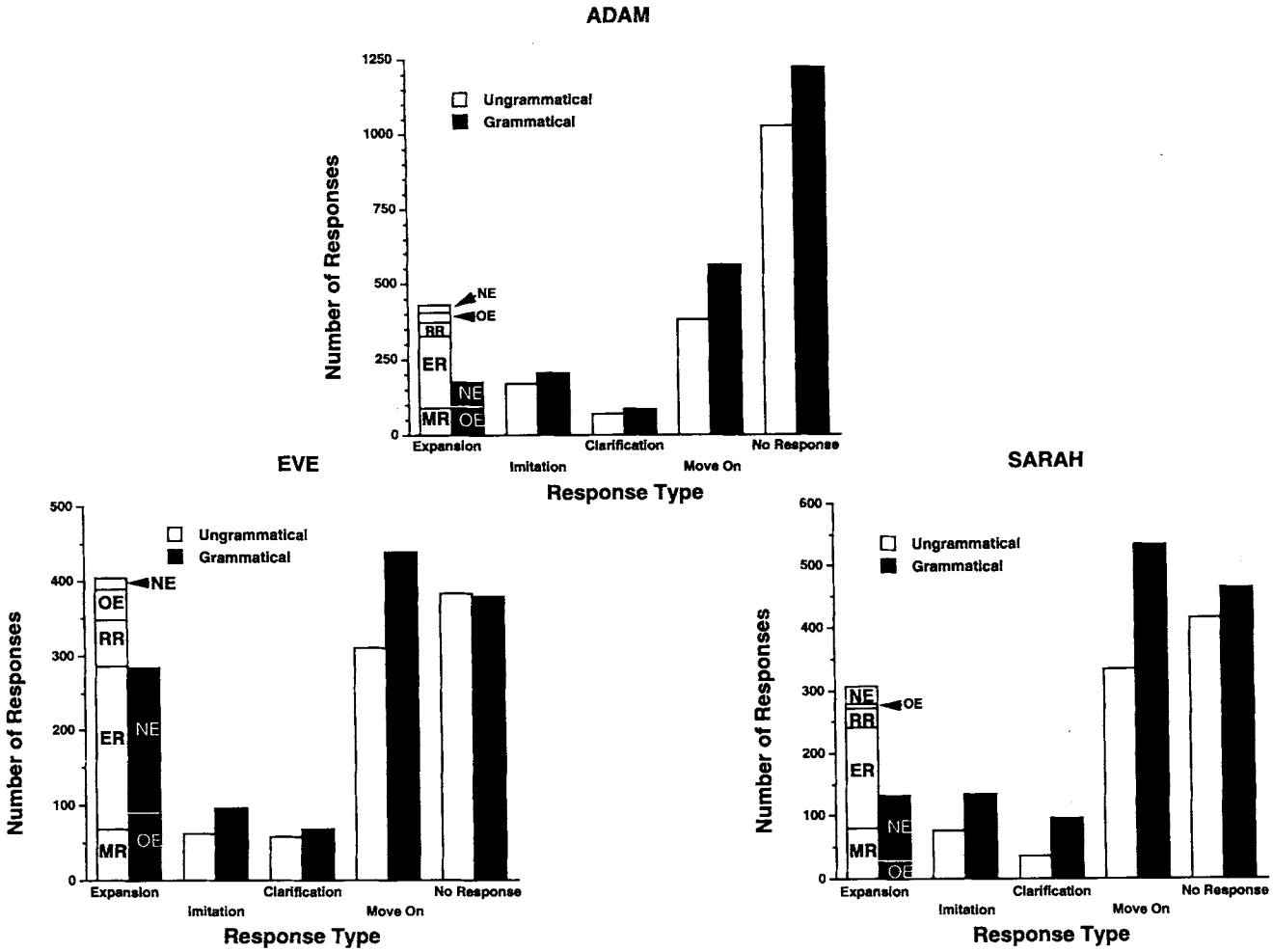

Figure 2. Frequencies of adult responses to grammatical article uses and ungrammatical article omissions. Subcategories of expansion: MR = minimal recast; ER = expanded recast; RR = related recast; OE = other recast; NE = nonrecast expansion.

same morphemes contained in other types of adult responses. However, these children are also three times more likely to repeat their original ungrammatical utterances following recasts than following other responses. Moreover, whereas children were somewhat more likely to repeat rather than amend their utterances following recasts, the opposite pattern occurred following other responses. Thus, although Farrar's study suggests that children may be able to discriminate recasts from other types of responses, it is mute on the question of whether recasts lead to the suppression of ungrammatical forms. Farrar acknowledged that the resolution of this question requires longitudinal data, of the sort we analyze here.

To assess the short-term effects of recasts, we conducted a series of 2×2 contingency table analyses asking whether grammatical noun uses were more likely to follow recasts of ungrammatical noun uses than move-on responses to ungrammatical noun uses. One set of analyses compared move-ons to target recasts (minimal recasts plus expanded recasts); another set compared move-ons to minimal recasts alone. We selected these contrasts because all previous investigations have agreed that move-ons (or "topic continuations") provide neutral feedback. These analyses were applied to three different time slices of our data—first to the entire body of data from each child, then to the data from each child following the point at which 50% spontaneous article usage was attained, and finally to the data from each child from each individual session depicted in Figure 1.

Our first set of analyses considered whether children were more likely to add articles from one noun use to the next depending on the response to the first use over the entire period of time we examined. We did not distinguish whether the child's next noun use immediately followed the adult response. For both Adam, $\chi^2(1, N = 692) = 11.8, p < .01$, and Eve, $\chi^2(1, N = 575) = 4.3, p < .05$, we did find associations between responses to ungrammatical noun uses and grammaticality of subsequent noun uses. Both children, however, used articles more often following move-on responses (Adam 34%, Eve 37%) than target recast responses (Adam 23%, Eve 29%). Analyses of noun uses following minimal recasts versus move-ons showed no significant associations.

Short-term effects of recasts might be more likely to be man-

ifest in connection with subsequent uses of the same noun that elicited the recasts. Our second set of analyses therefore considered whether children were more likely to add articles from one use of a given noun to the next depending on the type of response elicited by the first use. We did not find significant associations for any child in either the target recast or minimal recast analyses. In many instances, however, succeeding uses of the same noun were considerably separated; such separation may have diluted the effects of recasts. Therefore, we recomputed the earlier analyses, this time considering only successive noun uses that occurred within one or two conversational turns. Again, we did not find significant associations for any child in either the target recast or the minimal recast analyses.

One possible explanation for our failure to find effects of recasts is that most recasts occurred early in acquisition, when the children's overall grammatical use of articles was quite low. As Demetras et al. (1986) have suggested, feedback may be relevant only at those times when children are actively acquiring the constructions that are recast. Our inclusion of data from periods when children may not have been actively acquiring articles thus may have obscured relations between recasts and grammaticality (for further discussion of this issue, see Gleitman et al., 1984). Therefore, in our third set of analyses, we repeated the earlier series of contingency table analyses, this time including only data beginning at the point at which children attained 50% article use. In analyses with sufficient numbers of observations, we used chi-square; otherwise we used Fisher's exact tests. In only one instance did we find a significant association between response type and subsequent grammaticality. This occurred with respect to Adam's next noun use, $\chi^2(1, N = 296) = 6.1, p < .05$. However, Adam was more likely to shift to grammatical article use following a move-on response (80%) than a target recast response (58%).

Perhaps the relation between recasts and grammaticality is only evident at specific points in time. Therefore, narrowing our focus yet further, in our fourth set of analyses we assessed contingencies between target recasts and move-ons elicited by ungrammatical noun uses and grammaticality of the following noun uses for each individual session, using Fisher's exact tests. By chance alone we would expect 4 of the 79 test statistics calculated to be significant at the .05 level. This is precisely what we observed: significant contingencies in two transcripts each for Adam and Sarah, but none for Eve. In 2 of the 4 instances (one each for Adam and Sarah), grammatical noun uses occurred more often following move-ons than target recasts. In the other 2 instances, the opposite pattern obtained. Thus, over 4 sets of analyses, we failed to find evidence that systematic relations between recasts and immediately subsequent grammaticality ever hold.

Although recasts do not appear to affect children's immediate shifts toward grammatical usage, it is possible that they exert more subtle corrective effects. For example, given that children have shifted from ungrammatical to grammatical usage, it may be that they are more likely to persist in grammatical usage if their shift followed a recast than if the shift followed some other type of response. Thus, recasts might help promote stability of grammaticality more than do other types of responses.

To test this possibility, in our fifth set of analyses, we extracted examples in which the children had shifted to grammatical from ungrammatical usage following either target recasts or move-ons and tabulated the grammaticality of the child's next noun usage. In one analysis, we included all noun uses; in a second analysis, we included only those instances in which the same noun was used all three times. Chi-square analyses of the 2×2 contingency tables failed to reveal any significant associations. The largest test statistic we calculated was for Adam, including all noun uses: $\chi^2(1, N = 220) = 2.07, p > .05$. Recasts do not seem to affect the stability of grammatical article usage.

Finally, following suggestions by Marilyn Shatz (personal communication, April 18, 1991), we considered whether recasts have effects on the quantity or the quality of children's spontaneous revisions. For present purposes, revisions were defined as instances in which a child used the same noun two or more times with varying grammaticality in the same conversational turn. Each pair of successive uses in which grammaticality changed was counted as a revision. The subsets of revisions in which children changed from ungrammatical to grammatical usage were counted as self-corrections.

In our sixth set of analyses, we considered two questions. First, we asked whether target recasts were more likely than move-ons (following either grammatical or ungrammatical noun uses) to instigate revision of the next noun usage, without regard to the direction of the revision. For Adam, 4% of target recasts were followed by revisions, and 4% of all move-ons were followed by revisions. For Eve, 2% of target recasts were followed by revisions, and 3% of all move-ons were followed by revisions. For Sarah, 1% of target recasts were followed by revisions, and 1% of all move-ons were followed by revisions. In general, the incidence of revisions did not appear to be contingent on the type of immediately preceding adult response.

Next, we asked whether the conditional probability of self-corrections given revisions was greater following target recasts than move-ons (following ungrammatical noun uses only). For Adam, 25% of revisions following target recasts were self-corrections (3 of 12), whereas 50% of revisions following move-ons were self-corrections (11 of 22). For Eve, 3 of 5 revisions following target recasts were self-corrections, whereas 3 of 6 revisions following move-ons were self-corrections. Sarah had no self-corrections following either target recasts or move-ons. Thus, we failed to find any evidence that recasts prompt self-corrections.

In summary, across six sets of analyses examining effects of individual recasts on children's shifts to grammatical usage, children's stability of grammatical usage, and children's self-corrections, we found no evidence to support the contention that recasts provide negative evidence and serve as corrections. To the contrary, the small number of significant relations that we did find predominately suggested that children may have been less likely to shift from article omission to article usage following recasts than following neutral adult responses.

Long-term effects of recasts. The most obvious feature of the data shown in Figure 1 is that the children's grammaticality increased across time. A significant linear relationship between proportional grammaticality and time existed for each child: Adam, $r = .84$; Eve, $r = .89$; and Sarah, $r = .90$; all $p < .01$. The cumulative number of recasts also has a linear relation with time. A plausible model for the effect of recasts might suggest that each recast incrementally strengthens the child's grammat-

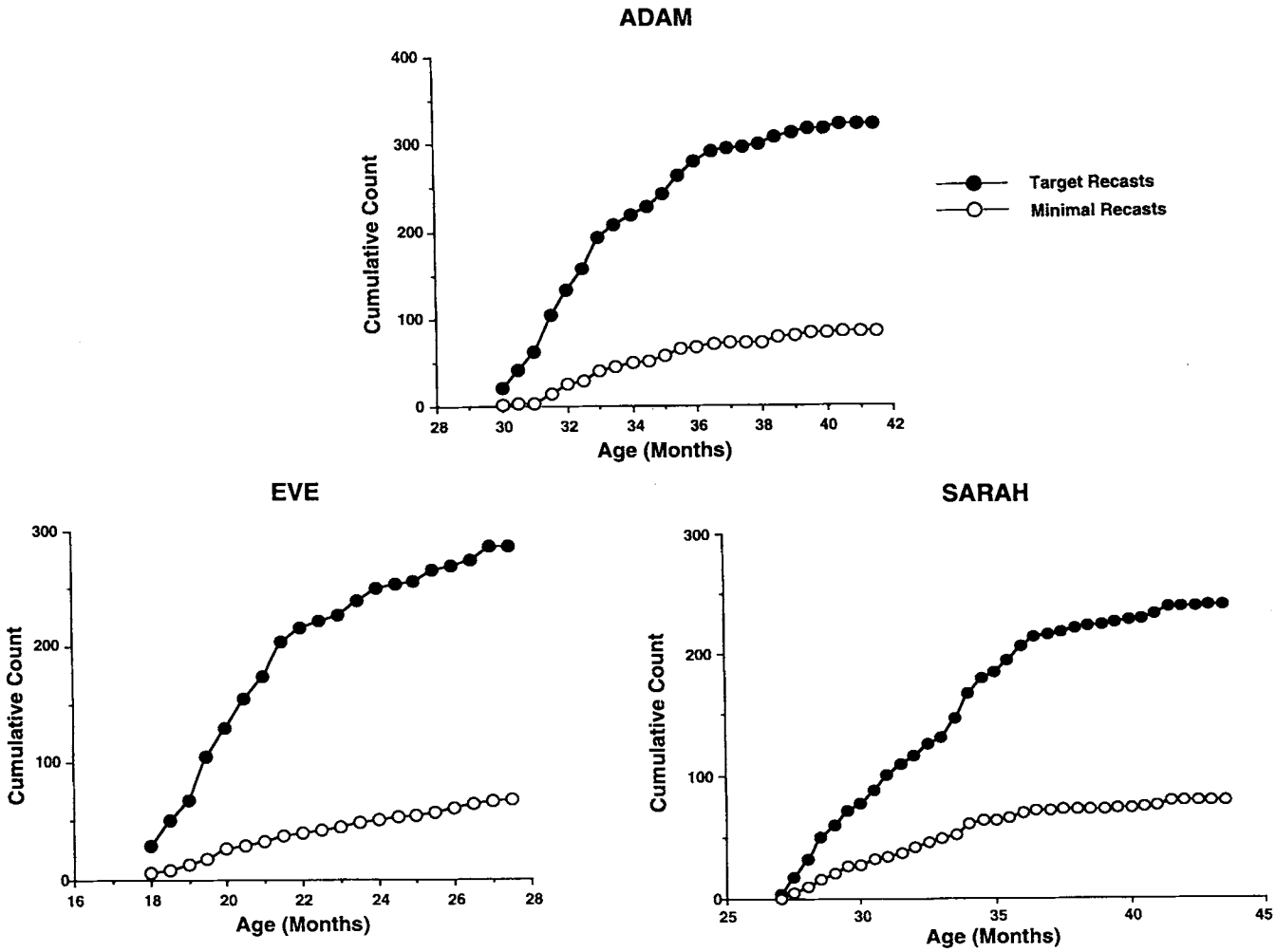

Figure 3. Cumulative frequencies of recasts of children's ungrammatical article omissions.

ical response. Under such a model, we would expect to find a positive linear relationship between proportional grammaticality and cumulative number of recasts. This is precisely what we found for all 3 children with both minimal recasts (Adam, r = .71; Eve, r = .87; and Sarah, r = .80) and target recasts (Adam, r = .68; Eve, r = .83; and Sarah, r = .83; all $ps < .01$). Cumulative numbers of recasts are shown for each child in Figure 3.

It is unclear how to interpret such correlations. After all, given that grammaticality increases across time, it would not be surprising to find a significant correlation between proportional grammaticality and any variable measured cumulatively, including, say, the number of animal crackers consumed by the child. In this last instance, we would certainly reject the correlation as being spurious. Given the model proposed earlier, however, we should expect proportional grammaticality and cumulative recasts to have a significant association over and above their common relationship with time. On this model, at times when recasts mount up more rapidly or less rapidly than usual, grammaticality should increase at correspondingly more rapid or less rapid rates. We would not expect to see this same sort of relation between grammaticality and cookies, where no causal link possibly exists. Therefore, we computed partial correlations for each child between proportional grammaticality at time T and cumulative recasts (either minimal or target) up to time T, controlling for time. We found no significant positive correlations. We did, however, find several significant negative correlations: for Adam's target recasts, $r = -.50$, $p < .05$; for Adam's minimal recasts, $r = -.55$, $p < .01$; and for Sarah's minimal recasts, $r = -.40$, $p < .05$. These results show that the simple incremental strengthening model proposed earlier cannot be correct.

It is possible, indeed even likely, that influences on children's grammars are not immediately manifest in their productions. Effects of recasts might be distributed over time, appearing gradually over several lags. The simple correlational approach used earlier is incapable of capturing this more subtle type of association. Moreover, correlations between variables measured contemporaneously may reflect causal relations flowing in either direction. Here, we are interested only in capturing possibly causal effects of recasts on grammaticality; possible effects of grammaticality on recasts can be filtered out by considering appropriately time-lagged correlations.

What we wish to explore is the possibility that the incidence of recasts may be a "leading indicator" of proportional grammaticality in the child, in much the same fashion that an economist might examine whether interest rates are a leading indicator of employment. Econometrics has evolved a set of time series analytic techniques that can be used to determine whether this sort of relation exists between two variables. We apply these techniques to examine possible relations between language input and acquisition. Because time series analyses, to our knowledge, have not been applied in this connection before, we provide a brief overview for readers unfamiliar with these techniques. Gottman (1981) offers an accessible, comprehensive introduction to time series analysis.

A data set comprising multiple measurements of a variable collected at equal time intervals constitutes a time series. The values of data points in a given time series may be in part determined by time itself (or by additional variables acting indirectly through time). For example, employment increases across time, as population increases. Employment also has regular seasonal fluctuations (e.g., increases in construction employment in the summer, decreases in the winter). Time series analysis begins by removing these deterministic components from the data, usually either by differencing or by regressing the raw data on control variables and retaining the residuals. A time series from which deterministic components have been removed is called *stationary.*

Given a stationary time series, one may ask whether future values of the time series are predictable on the basis of some combination of past values of the series. One way to address this univariate question is to compute a multiple regression, using lagged values of the series as predictors (e.g., one possible predictor would include values of the time series shifted by one time period [a single lag]; a second would include values shifted by two time periods [two lags], and so forth). Because this analysis involves regressing a time series on itself, it is called an *autoregression.*

Given two stationary time series, one may ask whether future values of one time series are more predictable on the basis of past values of both time series than they are on the basis of past values of only the series being predicted. If so, then one variable is a leading indicator of the other. For example, if future employment is more predictable given information about both past interest rates and employment than information about past employment rates alone, then interest rates are a leading indicator of employment. Here, we refer to terms in the regression model drawn from lagged values of the leading indicator time series as *cross-regressive.*

To return to the question at hand, we explored the possibility that recasts are a leading indicator of grammaticality by determining whether a mixed model including both autoregressive terms providing weighted measures of the child's past grammaticality and cross-regressive terms providing weighted measures of past input predicted the child's current behavior better than a purely autoregressive model. If recasts do affect overall grammaticality, we would expect the mixed models to be significantly better predictors. Moreover, if recasts serve as corrections, we would expect significant cross-regressive terms to be positively weighted: Higher levels of recasts at early points in time should be associated with higher levels of grammaticality at later points in time.

For the input time series, we calculated the numbers of minimal recasts and target recasts per child noun use in each transcript. These measures index the general frequency of minimal and target recasts. For the predicted time series, we used measures of the children's proportional grammaticality. To ensure that the time series were stationary, we factored out significant linear trends. The resulting time series (for minimal recasts) are displayed in standardized form in Figure 4.

We used a four-step procedure outlined by Gottman and Ringland (1981) to assess leading indicator effects of recasts, as shown here:

1. Fit a large regression model including more auto- and cross-regressive terms than necessary to the data.
2. Iteratively remove the highest order terms (representing the longest lags) from the regression model so long as they fail to meet a criterion of significance.
3. Fit a model including only the autoregressive terms from the final model in Step 2 to the data.
4. Compare the goodness of fit of the models constructed in Steps 2 and 3 by calculating the maximum likelihood statistic,

$$Q = D * \ln(SSE_3/SSE_2),$$

where D is the number of points predicted in the models and SSE_2 and SSE_3 are the residual sums of squares in Models 2 and 3, respectively. For large samples, Q is distributed as a chi-square with degrees of freedom equal to the difference in the number of parameters in the two models (i.e., the number of cross-regressive terms in Model 2).

Although time series analytic methods are typically applied to data sets larger than those available to us, with suitable cautions it is possible to apply these methods to smaller data sets as well (Jones, 1991). We grappled with two problems arising from our small sample sizes. First, Q approximates chi-square for large samples only. Therefore, for each analysis reported here, we estimated the distribution of Q by conducting Monte Carlo simulations, each using 1,000 pairs of time series randomly drawn from normal distributions with the same means and variances as those observed in our data. Second, the confidence intervals around terms in the regression equations are inversely proportional to sample size. Thus, small samples provide little power for discerning significant terms. To avoid Type II errors, we used a liberal criterion for retaining regression terms, $p < .20$ (for larger samples, Gottman and Ringland [1981] advocated using $p < .10$). This strategy, however, increases the risk of Type I errors in our analyses. In the present circumstances, therefore, finding corroborating evidence across different children is critical.

Our choice of the number of lags to use in constructing the initial models for each child was constrained by two factors. First, given results of Baker and Nelson (1984) and Nelson (1977) showing short-term facilitative effects of recasts, it was our intuition that corrective effects of recasts ought to appear over reasonably short periods of time—certainly no longer than 3 months. Second, the first several data points in each predicted time series had to be discarded because lagged values for one or more predictors were missing. In the present instance, this was not totally without benefit, because several authors have sug-

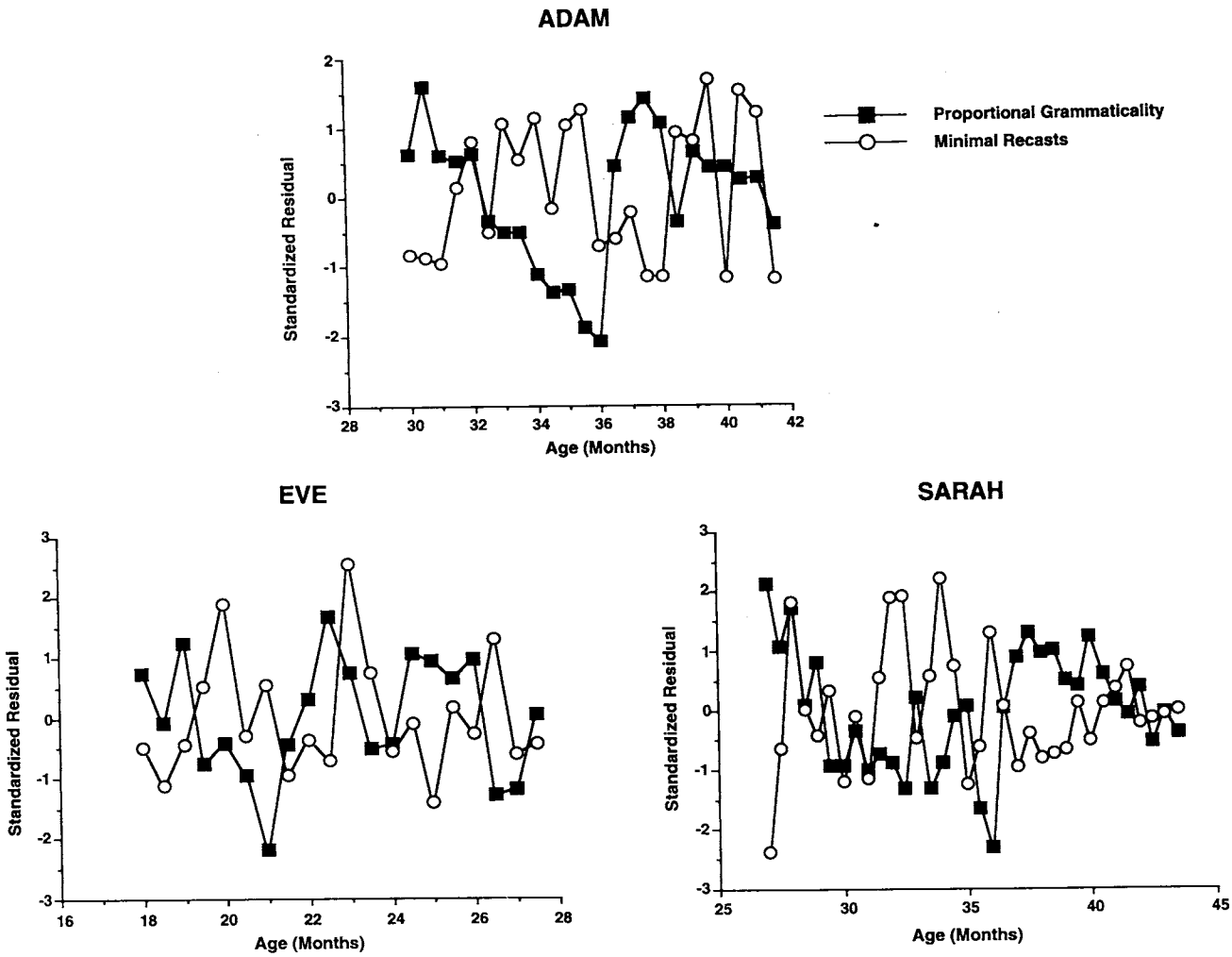

Figure 4. Transformed (stationary) time series for proportional grammaticality and minimal recasts.

gested that recasts (or other corrections) may not have effects until children are "ready." Discarding the earliest data points allowed us to focus on those periods in which the children were more likely to have attained the appropriate state of readiness. On the other hand, we wished to minimize data loss. We therefore decided to begin with four lagged terms for Eve (up to 8 weeks), with five lagged terms for Adam (up to 10 weeks) and with six lagged terms for Sarah (up to 12 weeks).

For all three pairs of time series using minimal recasts per noun use as the predictor, we found significant differences in the goodness of fit between the mixed and purely autoregressive models. For Adam, $Q(2) = 10.11$, $p < .05$; for Eve, $Q(3) = 10.04$, $p < .05$; and for Sarah, $Q(5) = 27.77$, $p < .01$. With target recasts per noun use as the predictor, we found a significant difference only for Sarah, $Q(3) = 10.95$, $p < .05$. Thus, we have corroborating evidence that information about minimal recasts in input did improve our ability to predict children's later grammaticality: Minimal recasts are a leading indicator of grammaticality.

The hypothesis that recasts serve as corrections entails that they be positive leading indicators of grammaticality: the more often recasts occur, the faster children's grammaticality should improve. The standardized weights of significant cross-regressive terms for each of the minimal recast analyses are shown in Table 2. Note that almost all of these weights (except for the

Table 2
Time Series Analyses of Children's Singular Count Noun Uses: Standardized Cross-Regressive Weights

	Minimal recasts		
Lag (weeks)	Adam	Eve	Sarah
2	−0.65****	−0.35*	−0.36***
4	−0.32*	−0.34*	
6		0.37*	−0.31**
8			−0.25*
10			0.57****

Note. Only coefficients for cross-regression terms remaining in the final equation (i.e., significant terms) are shown.
* $p < .20$. ** $p < .10$. *** $p < .05$. **** $p < .01$.

longest lags) are negative. Thus, all 3 children provide evidence corroborating that recasts are negative leading indicators of grammaticality. This is the opposite of what would be expected if recasts were serving as corrections.

Summary

Several previous studies (Hirsh-Pasek et al., 1984; Morgan & Travis, 1989; Penner, 1987) have shown that recasts are most strongly contingent on ungrammaticality when children are in the third year of life. During this time, children usually begin to acquire articles. Here, we discovered that children's article omissions are recast at exceptionally high rates. These facts make article use an ideal case for the study of possible corrective effects of recasts.

To examine these effects, we first conducted a variety of contingency table analyses of effects of recasts on both subsequent grammaticality and self-correction. These generally failed to reveal any influence of recasts as compared with move-ons. We next conducted time series analyses of the cumulative effects of recasts on overall grammaticality. Across all 3 children, these consistently revealed that recast are negative leading indicators of children's grammatical usage of articles.

These results are incompatible with the hypothesis that recasts serve as corrections. If recasts are causally linked to acquisition of syntactic constructions, they must serve alternative function. We suggest below that the role of recasts may primarily involve fostering children's grammatical diversity.

Study 2

To ensure that our results do not simply reflect idiosyncratic properties of article acquisition, we explored the role of recasts in children's acquisition of a second syntactic phenomenon: auxiliary inversion in Wh questions. Children use Wh questions commonly, often using ungrammatical forms. Thus, the study of Wh questions affords frequent opportunities to observe effects of recasts.

In English, the most common type of Wh question features a preposed Wh constituent (what, who, where, when, how, why, what kind of *X*, which color, etc.) that must be immediately followed by an auxiliary verb carrying the tense of the matrix sentence. Wh questions in other languages may lack one or both of these features. In some languages, Wh constituents remain in their original position (as they do in English occasional questions such as "You want to eat what?"). In other languages, the Wh constituent appears at the beginning of the sentence, but no subject-auxiliary inversion occurs.

The development of Wh questions in English-speaking children has been described by Brown (1968), Brown, Cazden, and Bellugi (1969), and Kuczaj and Brannick (1979), among others. When Wh questions first appear in children's productions, they are generally lacking auxiliary verbs. Some children later go through a period in which they produce Wh questions with auxiliaries in noninverted position. Here, we were concerned with children's observation of the syntactically obligatory inversion of auxiliaries in questions with preposed Wh constituents. As in Study 1, our fundamental question was whether parental recasts of children's Wh question auxiliary inversion errors are associated with lower future levels of such errors.

Table 3
Types of Responses to Children's "Wh" Questions

Category	Child utterance	Adult response
Minimal recast	What dat come from?	What did that come from?
Expanded recast	Where de light, Mommy?	I don't know. Where did the light go?
Other recast	Where is birthday cake?	Where is the birthday cake? I don't know.
Nonrecast expansion	Old Lady, what do you like to read?	Old Lady, what would you like to read?
Imitation	How do you know?	How do I know? Oh, I know everything.
Clarification question	Why dat has to retach to dat?	Why what?
Confirmation question	Pap, where you cookbook?	My cookbook?
Move-on	Why you didn't put the top on?	Because there's too much rubbish in it, that's why. Any other questions?

Note. *Wh* questions include who, what, when, where, how, and why.

Method

Adam's, Eve's, and Sarah's use of inverted auxiliary verbs in Wh questions was tracked from the point when such use began to the end of each child's transcript set. A computer search established the initial point. Included in our tabulations were transcripts from Adam between the ages of 2 years 9 months to 4 years 10 months (Transcripts 14 to 55), Eve from 1 year 11 months to 2 years 4 months (Transcripts 11 to 20), and Sarah from 2 years 9 months to 5 years 1 month (Transcripts 30 to 139).

As in Study 1, two exhaustive files were prepared. The first, "ungrammatical" file was built by a computer search that identified all instances of utterance-initial Wh words that were not immediately followed by an auxiliary. The second, "grammatical" file was built by a computer search that identified all instances of utterance-initial Wh words that were followed by an auxiliary somewhere in the same utterance. Entries in each file consisted of the entire conversational turn containing the target use or uses and the following adult conversational turn, with notation of the location in the transcript of each turn.

As before, each file was coded in two passes, the first to decide whether each possible Wh question should be retained for further coding, and the second to assign the adult response to each grammatical or ungrammatical Wh question to one of nine categories. These categories were identical to those used in Study 1, except that the *related recast* category was not relevant here. Table 3 provides examples of each response type.

J.L.M. and L.L.T. shared responsibility for coding the adult responses, each coding about half of the transcripts. Approximately one quarter of each coder's half was blind coded by the other coder for reliability. Agreement was greater than 90% for all transcripts.

Results and Discussion

Our data set included more than 4,000 Wh questions. We found slightly more ungrammatical Wh questions for Adam and Eve but more grammatical Wh questions for Sarah. Adam's data included 1,296 grammatical and 1,553 ungrammatical Wh questions, Eve's data included 131 grammatical and 247

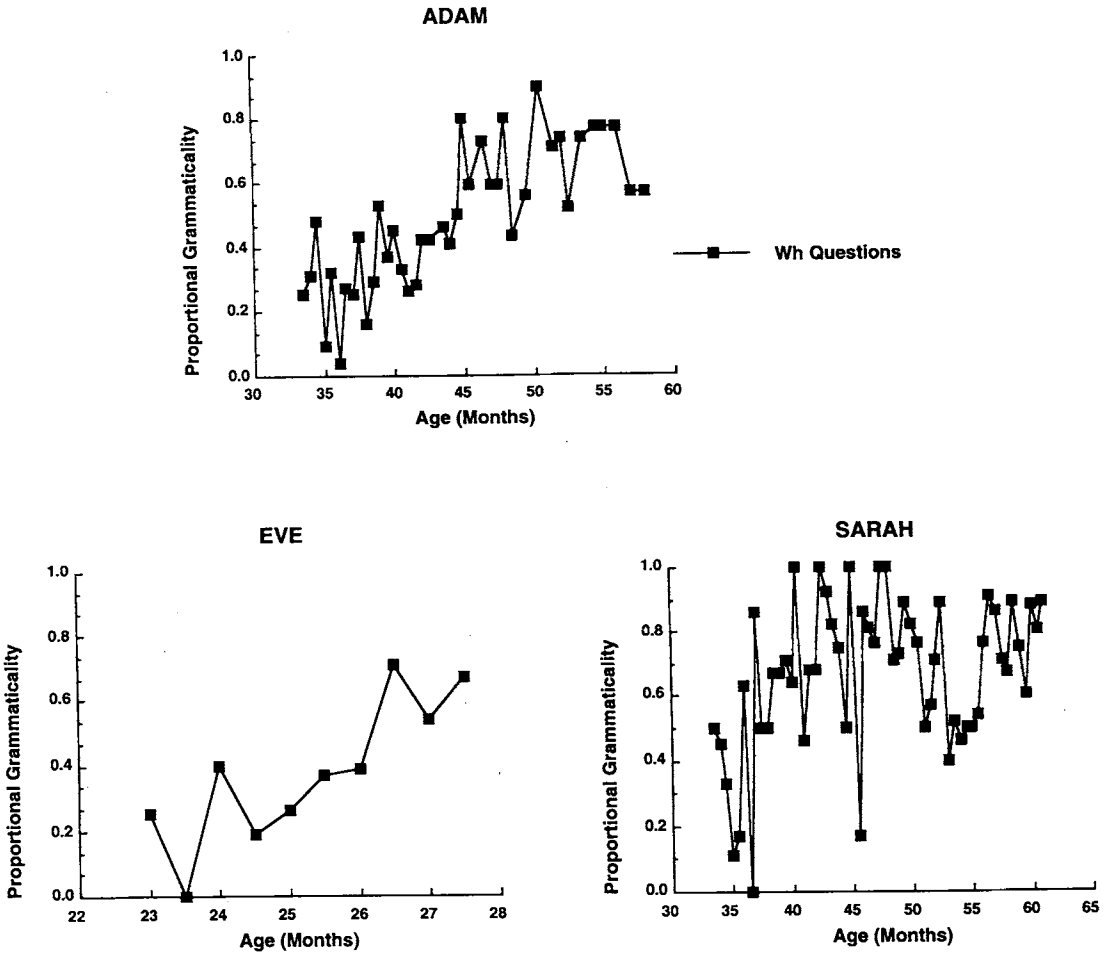

Figure 5. Proportions of children's grammatical *Wh* (who, what, when, where, how, and why) questions.

ungrammatical Wh questions, and Sarah's data included 544 grammatical and 269 ungrammatical Wh questions. The proportion of grammatical Wh questions in each transcript is shown for each of the 3 children in Figure 5. As in Study 1, successive pairs of Sarah's transcripts were combined.

Frequencies of responses to children's grammatical and ungrammatical Wh questions are shown in Figure 6. Because the number of confirmation questions was so low, these were combined with clarification questions. Chi-square analyses of these data (combining all types of recasts and nonrecast expansions in the superordinate category *expansion*) showed significant associations between response type distribution and grammaticality for all 3 children: Adam, $\chi^2(4, N = 2{,}849) = 67.3$, $p < .01$; Eve, $\chi^2(4, N = 378) = 65.6$, $p < .01$; and Sarah, $\chi^2(4, N = 813) = 27.3$, $p < .01$. For all 3 children, expansions were more likely to occur in response to ungrammatical Wh questions, whereas imitations and move-ons were more likely to occur in response to grammatical Wh questions. Again, this pattern of results replicates that found in previous studies of differential adult responses to grammatical and ungrammatical child utterances.

In Eve's data, target recasts formed a large proportion of actual adult responses to ungrammatical Wh questions (26%). In Adam's and Sarah's data, however, they formed a small proportion of such responses (8% and 6%, respectively). Because the proportions and numbers of recasts for each child were much lower than in Study 1, we carried out only a subset of the analyses pursued there.

Short-term effects of recasts. As in Study 1, we conducted 2 × 2 contingency table analyses asking whether grammatical Wh questions were more likely to follow recasts of, or move-on responses to, ungrammatical Wh questions. We computed four sets of analyses for each child: Two examined effects of minimal recasts; the others examined effects of target recasts. Crossed with this, two analyses examined effects on the grammatical status of the next Wh question, whereas the others examined effects on the next Wh question of the same type that elicited the response. In no instance did we find a significant association between the type of response and the grammaticality of the succeeding Wh question. The largest coefficient we computed was for Adam's target recasts–same type question, $\chi^2(1, N = 365) = 3.36$, $p < .10$. In this case, grammatical Wh questions occurred more often than expected following move-ons and less often than expected following target recasts.

Long-term effects of recasts. First, as in Study 1, a significant linear relationship between proportional grammaticality and

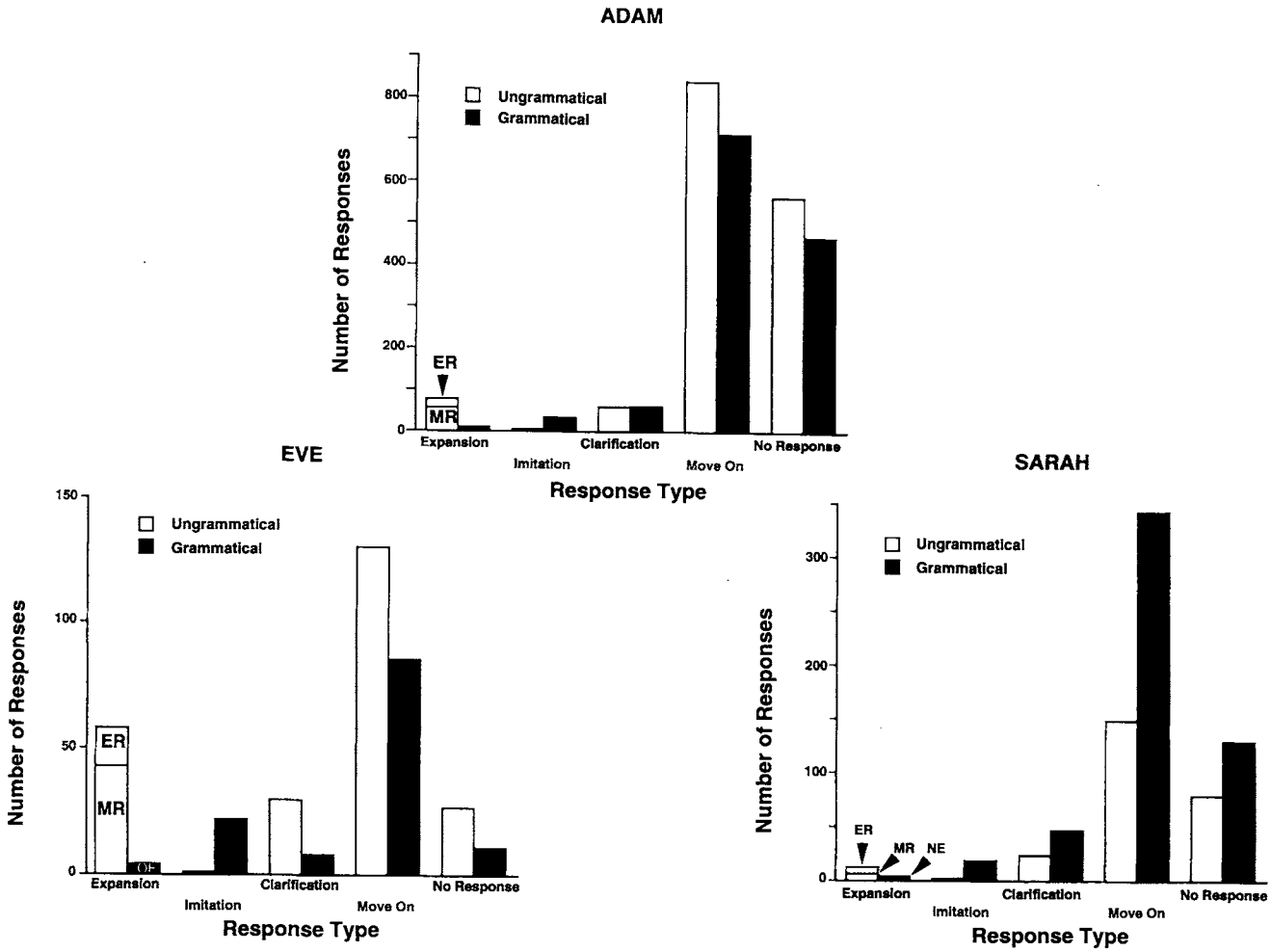

Figure 6. Frequencies of adult responses to grammatical and ungrammatical *Wh* (who, what, when, where, how, and why) questions. Subcategories of expansion: MR = minimal recast; ER = expanded recast; OE = other recast; NE = nonrecast expansion.

time existed for each child: Adam, $r = .80$; Eve, $r = .82$; and Sarah, $r = .35$; all $ps < .01$. The cumulative number of recasts for each child also had a linear relation with time. Cumulative numbers of recasts are shown for each child in Figure 7.

To test whether proportional grammaticality and cumulative number of recasts are linearly related, over and above their common relationship with time, we computed partial correlations for each child between proportional grammaticality at time T and cumulative recasts (either minimal or target) up to time T, controlling for time. As in Study 1, we found no significant positive correlations, although we did find significant negative correlations for Sarah's minimal recasts, $r = -.53$, $p < .001$, and Sarah's target recasts, $r = -.33$, $p < .01$. Thus, we again failed to find evidence in support of a simple incremental strengthening model.

Second, we conducted sets of time series analyses as in Study 1 on Adam's and Sarah's data. Too few data points were available for Eve to conduct meaningful analyses. As before, our expectations were that if recasts are leading indicators of grammaticality, then the larger mixed models should be significantly better predictors of children's grammaticality than the smaller, purely autoregressive models. Moreover, if recasts serve as corrections, then significant cross-regressive terms should be positively weighted.

As before, for the input time series, we calculated the numbers of minimal recasts and target recasts per child Wh question use in each transcript. Because the number of available data points per child was sufficiently large, we used the chi-square approximation to Q (rather than conducting additional Monte Carlo simulations), and we used the criterion ($p < .10$) suggested by Gottman and Ringland (1981) in deciding whether to retain terms in the regression equations. For Adam, we began with six lagged terms (up to 12 weeks). For Sarah, we began with seven lagged terms (up to 14 weeks).

For both children, the mixed models provided better predictions of grammaticality than did the corresponding purely autoregressive models. This was true by default for Adam, as the only significant terms in the regression equations for his data were cross-regressive. For Sarah, for both pairs of time series, we found significant differences in the goodness of fit be-

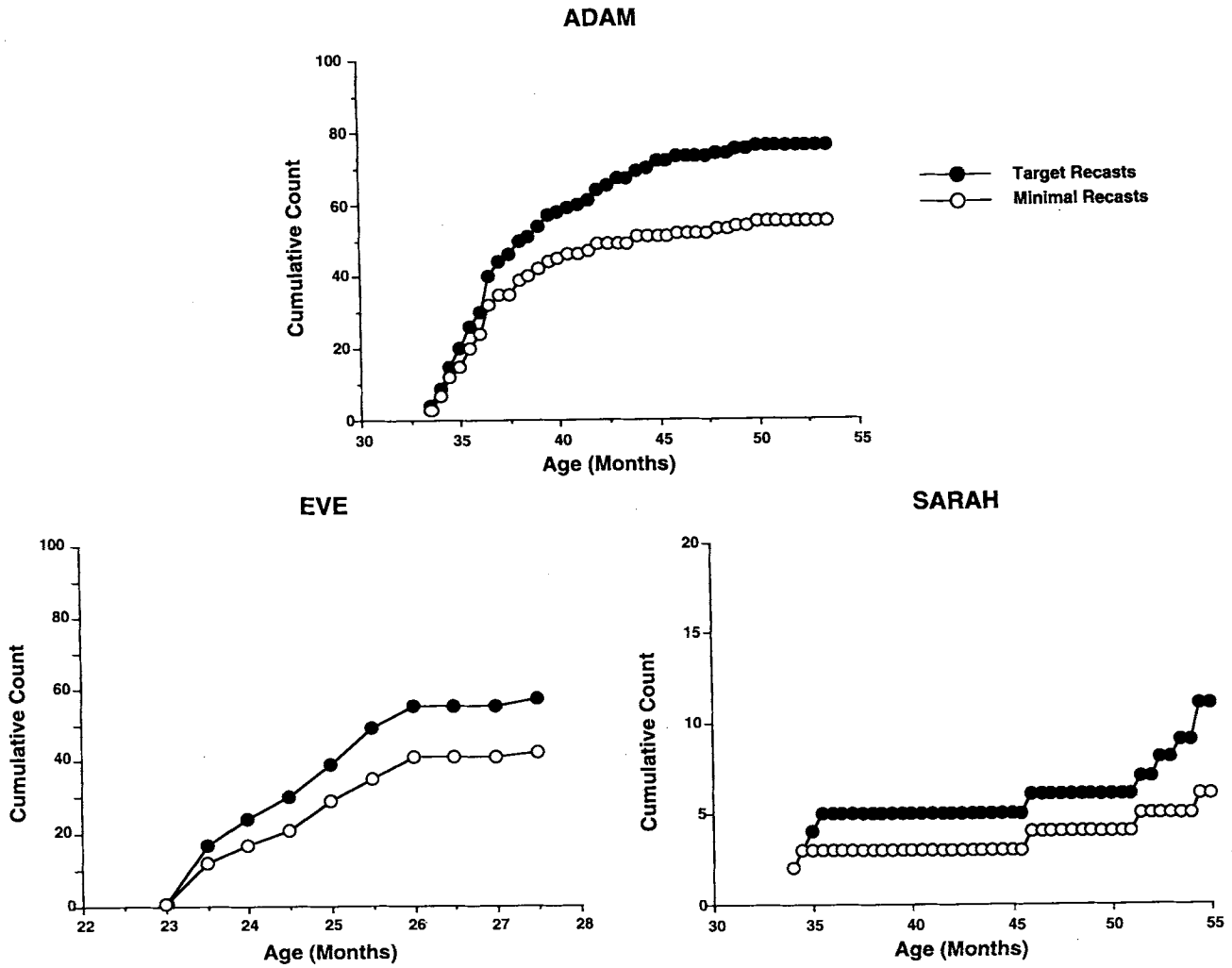

Figure 7. Cumulative frequencies of recasts of children's ungrammatical *Wh* (who, what, when, where, how, and why) questions.

tween the mixed and purely autoregressive models: minimal recasts, $Q(1) = 9.24$, $p < .01$, and target recasts, $Q(6) = 27.99$, $p < .01$. As in Study 1, information about recasts in input did improve our ability to predict children's later grammaticality.

The standardized weights of significant cross-regressive terms for each of the analyses are shown in Table 4. As in Study 1, the weights of the cross-regressive terms, particularly those representing shorter lags, were predominately negative. Again, the relationship between recasts and grammaticality did not take the form that would be expected if recasts served as corrections.

We also wished to calculate sets of time series based on children's productions and adults' recasts of specific types of Wh questions. Sarah received too few recasts for her data to be partitioned in this manner. However, we did examine lagged effects of recasts on Adam's *what* and *where* questions, his most frequently used types (1,432 and 594 tokens, respectively). Both target recasts, $Q(3) = 11.56$, $p < .01$, and minimal recasts, $Q(3) = 12.29$, $p < .01$, were leading indicators of the grammaticality of Adam's *what* questions, although neither class of recasts was a leading indicator for *where* questions. Weights of cross-regressive terms for *what* questions are also shown in Table 4; consistent with all our other analyses, these were negative. Thus, considering either Wh questions en masse, or individual types of Wh questions, we found no evidence that recasts serve as corrections.

Summary

Given that we found evidence that recasts lack corrective effects with respect to children's acquisition of articles, it is not surprising that we again found such evidence with respect to acquisition of auxiliary inversion in Wh questions. During most of the period in which Wh question auxiliary inversion is being acquired, the contingencies between recasts and ungrammaticality are weak or nonexistent (Hirsh-Pasek et al., 1984; Morgan & Travis, 1989). Also, for the 2 older children, the frequencies of recasts were quite low. As in Study 1, our contingency table analyses failed to reveal any influence of individual recasts

Table 4
Time Series Analyses of Children's "Wh" Questions: Standardized Cross-Regressive Weights

	All Wh questions		What questions
Lag (weeks)	Adam	Sarah	Adam
	Minimal recast		
2		−0.41***	
4			
6	−0.34**		−0.49****
	Target recasts		
2		−0.66****	
4			
6	−0.35**		−0.50****
8			
10	0.34**		
12		0.37***	

Note. *Wh* questions include who, what, when, where, how, and why. Only coefficients for cross-regression terms remaining in the final equation (i.e., significant terms) are shown.
** $p < .10$. *** $p < .05$. **** $p < .01$.

on subsequent grammaticality. Also as in Study 1, our time series analyses of the cumulative effects of recasts on overall grammaticality revealed that recasts were negative leading indicators of grammaticality. Again, this pattern does not comport with the hypothesis that recasts serve as corrections.

General Discussion

In our longitudinal studies of the effects of recasts on children's acquisition of articles and Wh questions, across a variety of analyses, we found no indications that recasts serve as corrections, either individually or cumulatively. Given the large number of tokens available to us, our analyses of effects of individual recasts had considerable power for revealing effects on children's grammaticality. Nevertheless, these analyses by and large failed to reveal any contingencies between recasts and either subsequent grammaticality or self-corrections. In contrast, our time series analytic examinations of cumulative effects of recasts did reveal effects on children's overall grammaticality. However, the relationships our analyses uncovered were opposite that which would be expected if recasts served as corrections. If recasts do occupy a causal role in acquisition, they must serve some noncorrective function. Here, we suggest that this alternative function may be to encourage diversity of usage, and we show that this explanation knits together several previous findings.

First, however, we note some possible objections to our conclusion that recasts do not generally serve as corrections in acquisition. It is clear that we did not conduct analyses assessing all possible relations between recasts and grammatical growth. In particular, the analyses we conducted were limited to those presupposing uniform effects of recasts across acquisition. One can imagine any number of alternative possibilities, involving thresholds, catastrophes, phase shifts, alternate equilibrium states, and so forth: The list is quite endless. But there are no principled reasons and no currently available evidence indicating that any of these more complex alternatives should be adopted. Moreover, imputing specialized learning mechanisms to the child would only serve to move empiricist explanations of language acquisition closer to existing nativist explanations. At the present time, the evidence supporting the possibility that recasts serve as corrections is confined to the observation that children attain grammaticality at some point after they have been exposed to recasts. However, causal arguments based solely on temporal sequences are fallacious (*post hoc ergo propter hoc*).

One alternative model appears at first sight to be maximally simple: recasts cause grammatical epiphanies. It is certainly possible that individual recasts may on occasion lead children to insights that their grammars are in error. However, it is unclear why recasts in particular should be crucial for such insights. In principle, errors could be made manifest by comparisons between children's own utterances and closely related parental responses, but they also could be revealed by covert comparisons between parental utterances and how the children themselves would have said the same sentences. On this latter view, any positive exemplar could instigate comparable insight. Note that recasts provide mere fractions of the total input for any particular construction. In the cases we examined, about 5% of all adult article uses and 3% of all adult Wh questions occurred in recasts. Even supposing that comparisons involving nonrecast positive exemplars impose more processing load, their overwhelmingly greater frequency makes it likely that most insights into errors will not be sparked by recasts.

Once children realize they are in error, they then must determine the proper generalizations needed for repairing their errors. No recast provides information about why an error is an error; therefore, no single recast can provide the solution to the problem. Children must first guess the nature of the problem—whether it is syntactic, semantic, morphological, phonological, or pragmatic—and must then guess the generality of the required repair. Do all nouns require articles? All common nouns? All singular nouns? All count nouns? All nouns referring to concrete entities? Whereas children may abruptly realize the need to mend their errant grammatical ways, actually doing so will take time, as they cast about for the proper generalizations. If recasts are crucial in this, we would expect to find periods in which recasts were associated with successively closer approximations to grammaticality. In our studies, we failed to find any such associations. Therefore, we think it unlikely that recasts occupy a privileged role in advancing acquisition through epiphanies.

If recasts do not serve as corrections, what is their function? On the basis of our finding that recasts are negative leading indicators of grammaticality, we suggest that recasts (or, more generally, expansions) serve an alternative purpose, encouraging variety of usage rather than repressing particular forms. It is not difficult to see how this could come about. As we pointed out in the introduction, language input includes positive evidence that recasts are not corrections. Parents reformulate their own utterances on occasion, thereby demonstrating variants of constructions to their children. Parents may also reformulate their children's grammatical utterances, only to use the same

constructions as the child at later points. These sorts of reformulation may impart the message, "That was one way of saying it. Now here's another." As Marcus (1993) and Valian (in press) pointed out, recasts are indistinguishable from such reformulations unless one has a knowledge of the grammar. Young children lacking such knowledge are therefore likely to interpret recasts as merely illustrating additional ways in which sentences may be constructed. Our hypothesis permits a unified explanation of effects of recasts and other types of expansions on acquisition.

This hypothesis allows us to fit together a spectrum of findings. First, on this view, we would expect expansions (including recasts) to have their strongest effects early in acquisition when children's grammatical repertoires are limited, and we might therefore expect parents to use expansions most often when their children are young. The empirical record bears this out: Hirsh-Pasek et al. (1984) and Morgan and Travis (1989) have shown that the frequency of expansions (particularly recasts) declines sharply in speech to children past the third year of life. Second, we would expect recasts to be effective in inducing children to add new grammatical forms to their repertoires when they had previously been using only corresponding ungrammatical forms. Nelson et al. (1973), Nelson (1977), and Baker and Nelson (1984) have demonstrated such effects. Third, we would expect recasts to reinforce children's ungrammatical forms as legitimate alternatives. This was what we observed in the studies reported here. Fourth, as parents intuit that recasts are encouraging new forms but not discouraging old, undesirable forms, they might be expected to turn away from using recasts as responses to ungrammatical utterances. Hirsh-Pasek et al. (1984), Morgan and Travis (1989), and Penner (1987) have shown that recasts are contingent on grammaticality in speech to 2-year-olds but that this contingency is weaker or nonexistent in speech to older children. Note that on the hypothesis that recasts serve as corrections, almost all of these findings are inexplicable.

If recasts do not provide correction, perhaps some other adult response, verbal or nonverbal, does. Although we cannot categorically deny this possibility, we think it is important to point out that corrections, if they exist, will have much less impact on theorization in language acquisition than is commonly supposed. This is because corrections are not equivalent to negative evidence.[5]

Recent claims that corrections are equivalent to negative evidence and can therefore replace innate constraints on learning may be traced in large part to a misconstruction of how early results in learnability theory by Gold (1967) have applied in the formulation of subsequent theories of acquisition. Gold was concerned with explicating the conditions under which languages contained in the mathematically defined classes of the Chomsky hierarchy could be identified. One of Gold's key results was the demonstration that input including negative evidence supports much more powerful learning than does input without such evidence.

Gold (1967) considered two input schemes. In the first (text presentation), the learner received only positive evidence, that is, examples of sentences that are in the input language. In the second (informant presentation), the learner received both positive and negative evidence, that is, appropriately labeled examples of sentences that either are or are not in the input language. With only positive evidence, it is possible for the learner to make mistakes involving errors of overgeneralization that will forever prevent the learner from arriving at the correct grammar. In Gold's (1967) words, "The problem with text is that if you pick too large of a language, the text will never tell you that you are wrong" (p. 461). Thus, in Gold's paradigm, only the most restricted class of the Chomsky hierarchy (the finite languages) is learnable given positive evidence. This class does not include the natural languages. In contrast, negative evidence provides learners with a means of avoiding or escaping overgeneralization errors; therefore, a vastly less restricted class of languages (a superset of the natural languages) is learnable given both positive and negative evidence.

Gold's (1967) work was an exercise in mathematical linguistics, not an attempt to construct a realistic portrayal of language acquisition. His model envisioned the learner as traversing a series of grammars, testing each one as a whole against past and present input, and having unlimited, though finite, amounts of time and data to arrive at a correct grammar. Subsequent theories have attempted to provide somewhat more realistic models of acquisition. As a result of these modifications, however, the impact of negative evidence has been subtly altered.

One obvious artifice in Gold's (1967) model is that children do not hypothesize and reject grammars of a piece, but rather construct the systems of knowledge underlying their use of language bit by bit. The topology of the language learning process (and of learnability models following Gold's) is thus more complex than depicted in Gold's model. Learners must traverse an n-dimensional hypothesis space wherein neighboring grammars differ incrementally and follow a path leading to a target region encompassing grammars of, say, the variants of English.[6] Adoption of this more complex topology, however, does not affect Gold's basic results. With positive evidence, there is no way to recover from errors of overgeneralization: Some of the subpaths in hypothesis space are one-way only, and as a result, there are times when the learner "won't be able to get there from here." Again, negative evidence can head off errors in the first place, or negative evidence can provide a means for learners to recover from overgeneralizations, ensuring that all subpaths are two-way and allowing the learner to move freely in hypothesis space.

More important, unlike Gold's (1967) theoretical learner, children have only limited amounts of input and time within which to accomplish their task. A central focus in the development of learnability theory has therefore been the construction of models that permit efficient learning. As considerations of efficiency enter in, distinctions that were without consequence in Gold's model have become critical. In particular, we can distinguish two forms of negative evidence. The first is supplied in response to errors the learner makes; this we call *corrective*. The

[5] Gordon (1990) has previously argued for this point, on the grounds that corrections are probabilistic and finite. Here, we develop a complementary argument.

[6] Hypothesis space encompasses all hypotheses available under a given set of constraints, whether or not they are immediately accessible to the learner. To move from one point to another, the learner need not consider all of the possibilities but rather need only to choose one.

second is supplied independent of the commission of any error; this we will call *cautionary*. Gold's definition of negative evidence was equivalent to what we are calling cautionary here, but this distinction was not important for his original result. Suppose that all negative evidence were corrective. Gold's learner, operating without time pressure, could afford to make any number of errors and wait around for corrections: eventually (i.e., "in the limit"), the learner would chance on a correct grammar.

Children do not have the luxury of acquiring language in such a leisurely fashion. As considerations of time come into play, the functions of cautionary and corrective negative evidence diverge. As Osherson, Stob, and Weinstein (1986) noted, the key to ensuring more efficient learning is minimizing backtracking by preventing learners from making too many mistakes in the first place. Corrections, because of their post hoc nature, cannot assist in limiting hypothesis space. Cautionary negative evidence, in principle, could ward off learners from unproductive regions of hypothesis space. However, because children, unlike Gold's (1967) learner, have imperfect memories, it is unclear whether even cautionary negative evidence would suffice to limit hypothesis space adequately. This last point is moot, because no one claims that children receive cautionary negative evidence: Parents never say, "Now listen carefully, here's an example of a sentence that's not English . . ."

Some alternative way of constraining hypothesis space must be found to ensure learnability within limits on time and input. Empirical justification for such constraints is readily available: Children commit only an infinitesimal fraction of logically possible errors in language learning. Certain enrichments of input, such as the inclusion of structural information (Morgan, 1986), can serve to restrict hypothesis space to some extent. Otherwise, constraints attributable to properties of the mind must be invoked. In fact, this is the motivation for most constraints in recent accounts (Lightfoot, 1991; Morgan, 1986; Pinker, 1984; Wexler & Culicover, 1980). Such constraints are required not to ensure learnability per se, but rather learnability with limited numbers of errors and from simple data, thereby ensuring efficiency.

A comparison of the models developed in Hamburger and Wexler (1973) and Wexler and Culicover (1980) illustrates the trade-offs between constraints and efficiency. These models made the same assumptions about the form of input, the learning procedure, and the type of grammar being acquired. Wexler and Culicover demonstrated learnability from vastly simpler data than did Hamburger and Wexler, but at the cost of assuming more than five times as many constraints.[7] The additional constraints in Wexler and Culicover allowed learnability from simple input; the inclusion of corrective negative evidence would not offset these constraints.

Conclusion

The argument that recasts provide negative evidence, thereby offsetting the need to appeal to properties of mind in formulating theories of language acquisition, rests on the premises that recasts serve as corrections and that corrections are equivalent to negative evidence. We have shown that both of these premises are mistaken. In the first instance, we have done so by adducing evidence showing that recasts are negative leading indicators of children's grammaticality, a pattern wholly incompatible with the contention that recasts provide correction. In the second instance, we have illuminated the theoretical misconceptions that led to conflation of the constructs *correction* and *negative evidence* and shown that corrections cannot supplant internal constraints in limiting children's grammatical hypotheses. In tandem with observations documented elsewhere concerning both the relative paucity of errors in acquisition and the contrast between the child's brief, limited exposure to language and the complexity and abstractness of the grammatical system that the child acquires, these results argue strongly for the importance of properties of the mind in guiding the child's development of language.

[7] The proof in Wexler and Culicover (1980) incorporated a set of slightly more than 20 constraints and demonstrated learnability of the transformational component of an Aspects-style grammar from input containing sentences with two levels of embedding or fewer. The proof in Hamburger and Wexler (1973) included the equivalent of three constraints; Wexler and Culicover estimated that the input required contained sentences with 400,000 levels of embedding or fewer.

References

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry, 10,* 533–582.

Baker, N., & Nelson, K. E. (1984). Recasting and related conversational devices for triggering syntactic development in children. *First Language, 5,* 3–22.

Bohannon, J. N., & Stanowicz, L. (1988). The issue of negative evidence: Adult responses to children's language errors. *Developmental Psychology, 24,* 684–689.

Braine, M. D. S. (1971). On two types of models of the internalization of grammars. In D. I. Slobin (Ed.), *The ontogenesis of language* (pp. 153–186). New York: Academic Press.

Brown, R. (1968). The development of wh questions in child speech. *Journal of Verbal Learning and Verbal Behavior, 7,* 274–290.

Brown, R. (1973). *A first language.* Cambridge, MA: Harvard University Press.

Brown, R., Cazden, C., & Bellugi, U. (1969). The child's grammar from I to III. In J. P. Hill (Ed.), *Minnesota symposium on child psychology* (Vol. II, pp. 28–73). Minneapolis: University of Minnesota Press.

Brown, R., & Hanlon, C. (1970). Derivational complexity and the order of acquisition of child speech. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 11–53). New York: Wiley.

Chomsky, N. (1959). Review of B. F. Skinner [*Verbal behavior*]. *Language, 35,* 26–58.

Demetras, M. J., Post, K. N., & Snow, C. E. (1986). Feedback to first language learners: The role of repetitions and clarification questions. *Journal of Child Language, 13,* 275–292.

deVilliers J., & deVilliers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2,* 267–278.

Farrar, M. J. (1990). Discourse and the acquisition of grammatical morphemes. *Journal of Child Language, 17,* 607–624.

Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28,* 90–98.

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language, 11,* 43–79.

Gold, E. M. (1967). Language identification in the limit. *Information and Control, 10,* 447–474.

Gordon, P. (1990). Learnability and feedback. *Developmental Psychology, 26,* 215–218.

Gottman, J. M. (1981). *Time-series analysis: A comprehensive introduction for social scientists.* New York: Cambridge University Press.

Gottman, J. M., & Ringland, J. T. (1981). The analysis of dominance and bidirectionality in social development. *Child Development, 52,* 393–412.

Hamburger, H., & Wexler, K. (1973). Identifiability of a class of transformational grammars. In K. J. J. Hintikka, J. M. E. Moravcsik, & P. Suppes (Eds.), *Approaches to natural language* (pp. 153–166). Boston: D. Reidel.

Hirsh-Pasek, K., Treiman, R., & Schneiderman, M. (1984). Brown & Hanlon revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child Language, 11,* 81–88.

Hoff-Ginsberg, E. (1985). Some contributions of mothers' speech to their children's syntax growth. *Journal of Child Language, 12,* 367–385.

Jones, K. (1991). The application of time series methods to moderate span longitudinal data. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 75–87). Washington, DC: American Psychological Association.

Kucera, H., & Francis, N. (1967). *An analysis of present-day English.* Providence, RI: Brown University Press.

Kuczaj, S. A., II, & Brannick, N. (1979). Children's use of the *wh* question modal auxiliary placement rule. *Journal of Experimental Child Psychology, 28,* 43–67.

Lightfoot, D. (1991). *How to set parameters.* Cambridge, MA: MIT Press.

MacWhinney, B., & Snow, C. E. (1985). The child language data exchange system. *Journal of Child Language, 12,* 271–296.

Maratsos, M. P. (1976). *The use of definite and indefinite reference in young children: An experimental study of semantic acquisition.* New York: Cambridge University Press.

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

McNeill, D. (1966). Developmental psycholinguistics. In F. Smith & G. Miller (Eds.), *The genesis of language* (pp. 15–84). Cambridge, MA: MIT Press.

Morgan, J. L. (1986). *From simple input to complex grammar.* Cambridge, MA: MIT Press.

Morgan, J. L., & Travis, L. L. (1989). Limits on negative information in language input. *Journal of Child Language, 16,* 531–552.

Nelson, K. E. (1977). Facilitating children's syntax acquisition. *Developmental Psychology, 13,* 101–107.

Nelson, K. E., Carskaddon, G., & Bonvillian, J. D. (1973). Syntax acquisition: Impact of experimental variation in adult verbal interaction with the child. *Child Development, 44,* 497–504.

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 109–149). Cambridge, England: Cambridge University Press.

Osherson, D., Stob, M., & Weinstein, S. (1986). *Systems that learn.* Cambridge, MA: MIT Press.

Penner, S. G. (1987). Parental responses to grammatical and ungrammatical child utterances. *Child Development, 58,* 376-384.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pinker, S. (1989). *Learnability and cognition.* Cambridge, MA: MIT Press.

Scarborough, H., & Wycoff, J. (1986). Mother, I'd still rather do it myself: Some further noneffects of 'motherese.' *Journal of Child Language, 13,* 431–438.

Skinner, B. F. (1957). *Verbal behavior.* New York: Appleton-Century-Crofts.

Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development, 43,* 549–565.

Valian, V. (in press). *Parental replies: Linguistic status and didactic role.* Cambridge, MA: MIT Press.

Wexler, K., & Culicover, P. W. (1980). *Formal principles of language acquisition.* Cambridge, MA: MIT Press.

Received September 18, 1991
Revision received July 30, 1993
Accepted April 20, 1994 ■

Correction to van den Boom and Hoeksma

In the article "The Effect of Infant Irritability on Mother–Infant Interaction: A Growth-Curve Analysis" by Dymphna C. van den Boom and Jan B. Hoeksma (*Developmental Psychology,* 1994, Vol. 30, 581–590), a paragraph was taken from two paragraphs on pages 221–222 of the discussion section of an article by Marilyn L. Riese, "Temperament Stability Between the Neonatal Period and 24 Months," published in *Developmental Psychology,* 1987, Vol. 23, 216–222, without proper citation or quotation. The authors regret this error.