

# Joint Attention and Early Language

Michael Tomasello and Michael Jeffrey Farrar

Emory University

TOMASELLO, MICHAEL, and FARRAR, MICHAEL JEFFREY. *Joint Attention and Early Language*. CHILD DEVELOPMENT, 1986, 57, 1454–1463. This paper reports 2 studies that explore the role of joint attentional processes in the child's acquisition of language. In the first study, 24 children were videotaped at 15 and 21 months of age in naturalistic interaction with their mothers. Episodes of joint attentional focus between mother and child—for example, joint play with an object—were identified. Inside, as opposed to outside, these episodes both mothers and children produced more utterances, mothers used shorter sentences and more comments, and dyads engaged in longer conversations. Inside joint episodes maternal references to objects that were already the child's focus of attention were positively correlated with the child's vocabulary at 21 months, while object references that attempted to redirect the child's attention were negatively correlated. No measures from outside these episodes related to child language. In an experimental study, an adult attempted to teach novel words to 10 17-month-old children. Words referring to objects on which the child's attention was already focused were learned better than words presented in an attempt to redirect the child's attentional focus.

By the time children begin productive language use, they have already established with their caregivers a variety of social-communicative routines. Ninio and Bruner (1978) and Ratner and Bruner (1978) analyzed the structure of these routines and demonstrated how such nonlinguistic interactions “scaffold” the child's early language. In effect, these interactions provide the young child with a predictable referential context that makes both her and her mother's language immediately meaningful. In his theoretical work, Bruner (e.g., 1981, 1983, 1985) has stressed that the underlying mechanism at work in these mother-child “formats” is joint attention. Because young children do not possess adult devices—either linguistic or nonlinguistic—for establishing the joint attention necessary for communication, recurrent interactive episodes help the infant to determine adults' attentional focus and thus the intended referent of their language. In this way, formats support early communicative interactions and so facilitate the child's early language development (see also Bakeman & Adamson, 1984).

Tomasello and Todd (1983) provided the first direct evidence that individual differences in the ability of mother-child dyads to

establish and maintain a joint attentional focus are related to the child's subsequent language growth. They videotaped mother-child dyads in their homes with a set of novel toys at monthly intervals for a period of 6 months, beginning with the child's first birthday. The amount of time dyads spent in joint attentional episodes during the 6 months was positively related to the child's vocabulary size at the end of this period. Several lines of evidence, including cross-lagged correlations, supported the argument that these episodes facilitated the child's early language development. This finding was replicated in a study comparing singleton and twin children (Tomasello, Mannle, & Kruger, 1986), in which positive correlations were found between time in joint attention at 15 months of age and vocabulary size at 21 months of age for each group of children separately as well as for the sample as a whole.

A second finding of these studies was that directiveness on the part of mothers—either verbal or nonverbal attempts to direct the child's attention or behavior—was negatively related to the proportion of object labels in the child's vocabulary. Others have found a similar relationship (e.g., Della Corte, Benedict, & Klein, 1983; Nelson, 1973). Nel-

We would like to thank the mothers and the children who made this study possible. Thanks also to Lauren Rio, Nancy Hodge, and Sara Mannle for their help in data collection and coding. This research was presented to the Boston Child Language Conference on October 25, 1985, and supported in part by NIMH Grant no. 1 R03 MH38997-01A1. Address for correspondence: Michael Tomasello, Department of Psychology, Emory University, Atlanta, GA 30322. Jeffrey Farrar is now at the Department of Psychology, University of Denver, Denver, CO 80210.

[*Child Development*, 1986, 57, 1454–1463. © 1986 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920/86/5706-0011\$01.00]

son (1981) hypothesized that this relationship is due to the child's inferences about the functional significance of language based on the way people around him use it. If adults use language primarily to refer to and categorize the world (e.g., naming novel objects), the child will infer that this is its primary function and the acquisition of object labels will be very important. Conversely, if adults are constantly using language for social-regulative purposes (e.g., to greet, thank, exhort, prohibit), the child will infer that this is its primary function and the acquisition of object labels will be less important.

Tomasello and Todd (1983) offered a different interpretation. They argued that adult directiveness makes it more difficult for child and adult to establish a joint attentional focus. When the adult attempts to redirect the child's attention in referring to an object, if the child is to determine the intended referent she must shift her attention so as to coordinate with the adult's. On the other hand, when the adult's reference follows into the child's already established attentional focus, the child need not actively make such a determination; coordination of attention depends only on the adult's skill at determining the child's focus. These authors thus argue that adult directiveness has an effect not so much on the child's overall assessment of the function of language but rather on the learning conditions surrounding the acquisition of individual words.

In the current view, then, joint attention is important for early language both on the "macro" level of extended periods of adult-child interaction and on the "micro" level of adult-child attempts to coordinate a specific piece of language with a joint attentional focus on its intended referent. The current study was designed to investigate these two types of attentional process in more detail. Two studies were conducted. The first was based on naturalistic observation of children just beginning to learn language. In contrast to Tomasello and Todd (1983), who looked only at time spent in joint attentional episodes, the current study focused on the language that occurred in these episodes. On the macro level, it was hypothesized that mother-child linguistic interaction would be facilitated when the interactants were jointly focused on some aspect of the nonlinguistic context. It was thus expected that inside, as opposed to outside, episodes of joint attention mother-child dyads would talk more and carry on longer conversations. Further, it was expected that while children would be en-

couraged to use longer sentences, mothers would use shorter sentences in these episodes because the intensity of these interactions encourages mothers' best Child Directed Speech register. In addition, because joint focus on the nonlinguistic context provides a predetermined conversational topic, mothers were expected to use less directive language inside joint episodes. On the micro level, it was hypothesized that object labels presented in an attempt to follow into the child's attentional focus would facilitate the establishment of joint attention and thus be positively related to the child's use of object-names. Conversely, object labels presented in an attempt to redirect the child's attentional focus were expected to discourage joint attention and thus be negatively related to the child's lexical acquisition. These relationships were expected to be stronger inside than outside macro-level episodes of joint attention.

The second study was a lexical training study designed to provide experimental corroboration for the findings on the micro level, that is, for the relationship between directiveness and lexical acquisition. Children in their second year of life were presented with novel object words either in an attempt to redirect their attentional focus or, alternatively, in an attempt to follow into their current attentional focus. It was predicted that children would learn new words more easily when they were presented in the latter condition. If this were indeed the case, it would provide experimental corroboration that adult directiveness is associated with slower early lexical acquisition, and it would provide evidence for the operation of attentional factors in this process.

## Study 1

### *Method*

*Subjects.*—Twenty-four white, middle-class children—equal numbers of firstborns and later-borns, males and females—were recruited by personal contact from local day-care facilities. Children were all between 12 and 18 months of age at recruitment (mean age = 14.6 months) and, according to maternal report, had begun productive language use.

*Observational procedure.*—Each mother-child dyad was videotaped at home for a period of 15 min on two occasions, once when children were 15 months and once when they were 21 months of age. Dyads were provided with a set of novel toys and given no special instructions except to "Do what you normally

would do." A research assistant and camera-person were present at each session. Mothers were told at recruitment that we were interested in their child's language development, and they were instructed at that time to begin noting the child's normal language practices. At each of the taping sessions, mothers were interviewed about their child's use of language.

*Coding procedure.*—Each videotape was first coded for episodes of joint attentional focus. As defined by Tomasello and Todd (1983), these were episodes that met the following conditions: (1) they began with one member of the dyad initiating interaction with the other, (2) both members then visually focused on a single object or activity for a minimum of 3 sec. (either member could look away briefly during an extended interaction), and (3) at some point during the joint focus (possibly at initiation) the child directed some overt behavior toward the mother (especially a look to the face) as evidence that he was aware of their interaction, thus excluding mere onlooking. An example might be: the child hands the mother a spoon, looking to her face; she places it in a cup; he takes it out, mouths it, and puts it back in the cup, looking to the mother; they continue this until someone (usually the child) shifts attention. Had the child played with these objects alone, this would not have been a joint attentional episode even if the mother was visually focused on the objects throughout.

The language inside and outside these episodes was of interest. It was coded in two ways. First, each videotape was transcribed by a team of two research assistants. An independent assistant coded the transcripts for specific language measures and then, using the joint attention coding, tabulated data on the language measures separately for inside and outside the episodes of joint attentional focus. Language measures for both mother and child were: number of utterances and Mean Length of Utterance (MLU). For mothers only, the proportional distribution of utterances into comments, questions, and directives was also determined. For children only, the total number of words and object-labels per minute was also determined. In addition, two measures of the dyad's conversational behavior were of interest: number of conversations (a conversation was defined as adjacent utterances on a common topic) and mean number of child turns per conversation (as an indication of conversation length). The comparison of these measures inside and outside episodes of joint attentional focus constituted the macro level of analysis.

The second coding, on a more micro level, concerned attentional factors associated with maternal reference to objects. In a separate coding of the videotapes, an independent team of two coders established for each maternal reference to an object (in which the object word received some prosodic stress): (a) whether or not it was made in an attempt to follow into the child's ongoing attentional focus (i.e., visual), as opposed to an attempt to redirect her attention or behavior, (b) whether or not the mother gestured or provided some other nonverbal indication of her attentional focus while making the reference, and (c) whether or not the child actually visually focused on the object at the time of the object reference. In this way, each object reference was assigned one of eight unique patterns generated by a factorial combination of the three dichotomies. Each of these patterns was designated by a sequence of three "+" or "-" symbols, one for each of the three criteria used in their determination. For example, an object reference in which the adult followed into the child's attentional focus, gestured, and the child focused on the object successfully was designated by "+++". These data were then tabulated separately for inside and outside periods of joint attentional focus.

It is important to note that these two levels of analysis are independent, not only in the methodological sense that they were coded independently, but also conceptually. Though on the surface it would seem that a joint attentional focus on the macro level would automatically imply that the mother's object references would follow into the child's focus, this is not necessarily so. First, joint attentional episodes sometimes involve several objects (e.g., placing blocks in a bowl). If a mother directed the child's attention to one of these objects and the child was focused on another, then this was considered a directive inside a joint episode. Second, a mother could make an attempt within a joint episode to redirect her child's attention to outside objects. If the child did not attend, or attended only briefly and then returned to the object of joint focus, this also was counted as a directive within a joint episode. Conversely, it could also happen that a mother could follow into her child's focus when not in a joint attentional episode. If this did not result in an extended (3-sec) period of joint focus, this was counted as an attempt to follow into the child's attention outside a joint attentional episode.

The language interview used to assess the child's language development at 15 and 21 months of age was an adaptation of the

Bates (1979) interview which, in addition to utilizing spontaneously generated information, prompts the mother to provide examples of the child's language use by asking her about specific contexts in which children talk. (For example: What does she do when she wants food? Any special foods? What about when she wants her bottle? When she's in the high chair? At the refrigerator? In the store?) From this interview, a vocabulary list (including pat phrases) was compiled. Vocabulary size was computed, as well as the proportion of the child's lexical items that were object labels (i.e., general nominals as defined by Nelson, 1973). This latter measure was used in an attempt to capture language-acquisition style independently of sheer size of vocabulary.

Reliability was computed for each of the measures by having a second team of assistants code 20% of the subjects and compute the percentage of their agreement with the original coders. Reliabilities were as follows: judgments of joint attentional episodes (durations had to be within 3 sec) agreed at 84%; child language measures (including conversation) agreed at 88%–100%; maternal language measures (including types of object reference) agreed at 82%–100%.

**Results**

At both time periods, mother-child dyads spent about two-thirds of their interaction time inside joint attentional episodes and about one-third of their time outside these ep-

isodes. Because of this difference, all measures of frequency were divided by the appropriate measure of time to yield a "per-minute" frequency. All other measures were proportions of one language measure relative to another. Each mother and child language measure was analyzed with a 2 × 2 repeated-measures ANOVA, using joint attentional state (inside and outside) and child age (15 and 21 months) as independent variables.

*Child and dyad language.*—Table 1 presents means and standard deviations for all child and dyad language measures. All of these measures were higher inside than outside the joint attentional episodes; for four of the six measures the difference was statistically significant. Inside, as opposed to outside, joint attentional episodes children produced more: utterances per minute,  $F(1,23) = 11.72, p < .05$ ; words per minute,  $F(1,23) = 10.02, p < .05$ ; and words referring to objects per minute,  $F(1,23) = 17.16, p < .01$ . The child's average number of turns per conversation was higher inside as opposed to outside joint attentional episodes,  $F(1,18) = 16.01, p < .01$  (only 19 dyads had conversations). Child age produced several main effects and interacted with joint attentional state for several of these measures, as shown in Table 1. In each case of interaction, differences between the values inside and outside joint attentional episodes were greater when the child was 21 months of age.

TABLE 1

LANGUAGE MEASURES INSIDE AND OUTSIDE JOINT ATTENTIONAL EPISODES AT BOTH CHILD AGES

LANGUAGE MEASURES	15 MONTHS		21 MONTHS	
	Inside Joint Episodes	Outside Joint Episodes	Inside Joint Episodes	Outside Joint Episodes
<b>Child:</b>				
Utterances (per min) . . . .	1.0 (1.4)	.6 (.82)	3.7 (3.1)	1.6 (1.3) <sup>*ab</sup>
MLU . . . . .	1.2 (.24)	.9 (.83)	1.3 (.29)	1.1 (.42)
Words (per min) . . . . .	1.2 (.17)	.8 (1.0)	4.9 (4.7)	2.0 (1.1) <sup>*a</sup>
Object labels (per min) ..	.6 (.10)	.5 (.75)	1.8 (2.0)	.8 (1.6) <sup>**a</sup>
<b>Dyad:</b>				
Conversations (per min) .	.5 (.51)	.4 (.41)	.9 (.87)	.7 (.69) <sup>a</sup>
Average child turns . . . .	1.7 (.73)	1.0 (.86)	4.5 (2.2)	2.4 (1.7) <sup>**a</sup>
<b>Mother:</b>				
Utterances (per min) . . . .	16.9 (8.2)	8.6 (5.6)	12.1 (5.1)	9.1 (4.7) <sup>**</sup>
MLU . . . . .	3.9 (.67)	4.2 (.65)	4.4 (.51)	4.9 (.10) <sup>*a</sup>
% Comment . . . . .	.56 (.11)	.46 (.18)	.48 (.09)	.41 (.20) <sup>*a</sup>
% Question . . . . .	.29 (.12)	.32 (.15)	.36 (.11)	.46 (.20) <sup>*a</sup>
% Directive . . . . .	.15 (.07)	.22 (.15)	.16 (.08)	.13 (.10) <sup>b</sup>

\* Inside and outside joint episodes different,  $p < .05$ .

\*\* Inside and outside joint episodes different,  $p < .01$ .

<sup>a</sup> 15 months different from 21 months,  $p < .05$ .

<sup>b</sup> Interaction between joint episode and child age,  $p < .05$ .

## 1458 Child Development

Because of the relatively high variability on some of the child and dyad measures, subjects were also assessed on an individual basis. The pattern of results remained the same. Out of 24 children, 20 had higher values inside joint attentional episodes on the utterances per minute and words per minute measures ( $p < .01$ , sign test), and 16 children had more object labels ( $p < .08$ ). Seventeen of the 19 children who had conversations had higher values inside joint episodes on both the conversations per minute and the average length of conversation measures ( $p < .01$ , sign test).

*Maternal language.*—Table 1 also presents means and standard deviations for maternal language measures. Like their children, mothers produced more utterances per minute inside as opposed to outside joint attentional episodes,  $F(1,23) = 15.87$ ,  $p < .001$ . Their MLUs, however, were shorter inside than outside the episodes,  $F(1,23) = 3.90$ ,  $p < .05$ . Of the maternal utterances, a higher proportion were comments,  $F(1,23) = 6.61$ ,  $p < .05$ , and a lower proportion were questions,  $F(1,23) = 4.07$ ,  $p < .05$ , inside the joint attentional episodes. Proportion of directives showed a significant interaction of joint attentional state and child age,  $F(1,22) = 6.06$ ,  $p < .05$ , such that mothers produced proportionally more directives outside joint episodes at the 15-month child age only (Newman-Keuls). As shown in Table 1, child age produced two main effects: comments became proportionally less frequent over child age, while questions became more frequent. There were no differences in the number of object references per minute made by mothers or in the distribution of these into the

eight object reference types as a function of joint attention; these are therefore not presented in Table 1.

*Maternal language and child lexical development.*—General measures of the mothers' language (utterances per minute; MLU; and proportion of comments, questions, and directives) did not correlate with the child's vocabulary size or proportion of nominals at 15 or at 21 months. What did correlate were the eight types of object references. Table 2 presents these correlations, controlling for child age in months. The overall pattern is quite striking and very clear-cut. Nothing the mothers did outside joint attentional episodes correlated, either positively or negatively, with either measure of the child's lexical development. Inside the joint attentional episodes, on the other hand, three of the object reference types in which the mother followed into the child's attentional focus (+ + +, + + -, + - -) correlated positively with either the child's vocabulary size, proportion of nominals, or both. For the most part, directive object references were negatively related to the child's language. Especially important was the finding that directives not accompanied by gestures and to which the child did not attend (- - -) were negatively associated with the child's vocabulary size at 21 months. Interestingly, the one positive correlation for directives was when the mother was gesturing to the object and the child focused on it successfully (- + +).

To help determine the direction of causality in these correlations, cross-lagged panel correlations were performed for each of the eight object reference types with both child

TABLE 2

PARTIAL CORRELATIONS OF MATERNAL OBJECT REFERENCES AT 15 MONTHS WITH CHILD LANGUAGE MEASURES AT BOTH CHILD AGES AS A FUNCTION OF JOINT ATTENTIONAL EPISODE

TYPE OF OBJECT REFERENCE	INSIDE JOINT EPISODES				OUTSIDE JOINT EPISODES			
	Vocabulary Size		% Nominals		Vocabulary Size		% Nominals	
	15 Months	21 Months	15 Months	21 Months	15 Months	21 Months	15 Months	21 Months
--- .....	-.30	-.46*	-.12	-.17	-.17	-.06	-.10	-.07
-- + .....	-.27	-.23	-.06	-.13	-.07	-.09	-.01	-.34
- + - .....	-.25	-.14	-.24	-.03	-.23	-.08	-.18	.03
- + + .....	.46*	.26	.38	.22	-.06	-.11	.04	.17
+ - - .....	.54*	.37	.47*	.20	.10	-.17	.18	-.34
+ + - .....	.45*	.34	.27	.21	.05	.08	.03	.31
+ - + .....	.17	.08	.20	.13	.13	.28	.23	-.04
+ + + .....	.50*	.62*	.44*	.61*	-.11	-.04	.14	-.17

\*  $p < .05$ .

language measures, both inside and outside joint attentional episodes. If the “opposite” cross-lagged correlations—that is, maternal measures at 21 months with child language measures at 15 months—are similar to those reported above, then it is likely that the child’s language is influencing the object reference types as much as the reverse. However, of the 32 “opposite” cross-lagged correlations, only one was statistically significant: frequency of the + + + model inside joint episodes correlated with child vocabulary size at .55,  $p < .05$ .

### Discussion

There were three main findings in this study. The first was that during periods of joint attentional focus both mothers and children talked more, the dyad engaged in longer conversations, and mothers used shorter sentences and more comments. It is tempting to conclude from this that, as hypothesized, periods of joint attentional focus in some way scaffold early mother-child linguistic interaction. However, another plausible hypothesis is that the causality is in the opposite direction: the dyad’s ability to interact linguistically is a major factor in the establishment and maintenance of joint attentional episodes. There is undoubtedly some truth to this. However, it is not the case that language is a necessary condition for a joint attentional focus—virtually every dyad had some joint interactions with no language. Nor is it the case that language is sufficient for joint visual attention—all dyads had linguistic interactions outside of joint attentional episodes. Also, it is important to note that while the child’s linguistic competence increased across the two observation sessions, the time in joint interaction did not. Thus, for all of these reasons the causality could not flow exclusively from language to joint attention. The most plausible interpretation, then, is that the direction of influence is “transactive”: joint attentional episodes scaffold the prelinguistic child into language, which helps the child establish and maintain these episodes, which facilitates further linguistic interactions.

The second finding was that the types of object references mothers made inside the episodes of joint attentional focus were related to the child’s subsequent language development, whereas these same measures outside the joint episodes did not correlate. This is despite the fact that there was no systematic difference between the types of models given inside and outside joint attentional episodes. However, because periods of heightened linguistic activity for the child

corresponded to joint attentional episodes, it may be that children are more tuned in to maternal language when they themselves are speaking or when they are engaged in conversations of a certain length. Thus, again, language may be part of the cause as well as the effect. Again, it is probably best to think in transactive terms. In this case, the causal factor may best be conceived as periods of joint attention, which involve linguistic as well as nonlinguistic elements.

The third main finding concerned the specific relationships between object reference types and the child’s language. Inside the joint attentional episodes, three of the four object reference types that followed into the child’s attentional focus correlated positively with the child’s subsequent lexical development, whereas one of the directive types correlated negatively. It is interesting to note that the only directive type that correlated positively was the one in which the mother made her attentional focus clear by gesturing and in which the child focused successfully on the referent object (− + +). It is puzzling at first glance that the + + − and the + − − reference types correlated positively with the child’s subsequent vocabulary since, in these, the child was not focused on the object at the precise moment the name was provided. However, by definition of follow-in, in both of these reference types the child was focused on the object when the mother began her utterance. These two types thus indicate situations in which the child looked away from the object before its name was uttered. Many times this simply meant that the child looked to the mother’s face as she spoke and returned to a focus on the object soon thereafter. Though in some cases the child shifted his attention permanently, most often these reference types do represent an instance of joint attentional focus and thus should facilitate the child’s word learning.

Once again in this third finding, however, either direction of influence is possible. In contrast to the current hypothesis that the object reference type affects the child’s lexical acquisition, it is possible that the correlations are due to the child’s influence on the mother: linguistically competent children induced mothers to provide certain types of object references. However, the cross-lagged correlations argue against this interpretation. Frequency of the − − − object reference type at 15 months correlated negatively with child language at 21 months, but child language at 15 months did not correlate with object reference types at 21 months. This pattern indi-

cates that the direction of influence is most likely from the object reference types to the child's language. In the case of the +++ type, the child vocabulary measure correlated with the object reference type both within and between time-points, and so the direction of influence is unclear. However, the same panel analysis of the "% Nominals" measure of child language produced the pattern favoring the interpretation that it was the object reference type that influenced the child's language and not vice versa; that is, there was no correlation between child language at 15 months and object reference type at 21 months. Overall, then, the most plausible interpretation of the pattern of correlations in the current study is that object references that follow into the child's attentional focus facilitate lexical acquisition, especially of object labels.

## Study 2

In an attempt to provide experimental corroboration for the third finding of the correlational study—the relationship between object reference type and the child's lexical acquisition—a lexical training study was designed. The focus was on the general finding that object labels given as the adult was following into the child's attentional focus were positively related to lexical acquisition, while those given as directives were negatively related.

### Method

**Subjects.**—Ten middle-class children, six males and four females, were recruited by personal contact from local day-care facilities. Children were between 14 and 23 months of age at recruitment (mean age = 17.4) and attended day-care on a daily basis. As determined by a maternal interview, all children were producing at least several words.

**Procedure.**—After some initial "warm-up" visits to the classroom, two research assistants saw children individually in a quiet room at the day-care facility. One researcher trained and tested the child, while the other observed and recorded her behavior. Each child participated in four training sessions, two per week for 2 weeks, as well as a follow-up testing session 2 weeks after the final training session. One session lasted 15–20 min.

Each child was assigned four objects from a set chosen to be unfamiliar to children of this age (e.g., gauge, clip, bow, wrench, etc.). The child was assigned objects so that they matched her phonological preferences, as determined by the maternal interview at recruitment. For each child, each of her four

objects was then randomly assigned to one of two attentional strategy conditions (follow-in or direct), such that there were two objects in each condition. In the follow-in condition, the experimenter waited until the child was engaged with the target object (visual and tactile contact) and then addressed the child with a short sentence in which the object word was stressed. In the direct condition, the experimenter waited until the child was not engaged with any object and then held up the target object and addressed her with a short sentence in which the object word was stressed. Half of the sentences in each condition modeled the word in the middle of the sentence (e.g., "The *clip* is here") and half modeled it at the end of the sentence (e.g., "Here's the *clip*").

At the beginning of the first session the child was asked for the name of each of her four objects. None of the children produced the correct name of any object. In each session, children were given four trials for each object. Each name was modeled once (order was randomly selected for each child for each session from the list of 24 possible orders), and then the entire sequence was repeated three more times. Any language the child used during the session was recorded. Productions that were judged by both researchers to be instances of the modeled word (judged on phonological similarity and contextual appropriateness) were recorded and labeled as either imitative (if they occurred directly after a model) or spontaneous. All of these productions (including imitations) constituted the spontaneous production measure. After all models had been given in a session, two tests were administered. First, in the elicited production test the experimenter simply held up each object (in random order) and asked "What is this?" If the child failed to respond, she was asked two times more. Second, in the comprehension task the experimenter placed the four objects side by side and asked for each object in turn (in random order with each object replaced after each trial) by instructing the child to "Give me the \_\_\_\_\_" and holding out his hand. Again, if the child failed to respond, he was given two trials more. Two weeks following the final training session, the elicited production and comprehension tasks were given again (and any spontaneous productions were noted) in a short follow-up session in which there was no training.

### Results

Table 3 presents means and standard deviations for the three dependent measures as

TABLE 3

MEANS AND STANDARD DEVIATIONS OF CHILD PERFORMANCE MEASURES AS  
A FUNCTION OF TRAINING CONDITION (Summed across All 4 Sessions)

Child Performance Measure	Follow-in	Direct
Frequency of spontaneous production . . . .	.40 (.50)	.40 (.50)
Frequency of elicited production . . . . .	.10 (.30)	.50 (.50)
Percent comprehension . . . . .	50 (09)	32 (10)*
Percent comprehension (follow-up) . . . . .	64 (13)	36 (13)*

\* Conditions different,  $p < .05$ .

a function of attentional strategy training condition. Because there was no systematic effect of session, the values used for analysis and presented in Table 3 are the values obtained by summing across the four training sessions. There was no effect of the placement of the word in the sentence (middle or end) or of the order of conditions. These were therefore excluded from further analysis.

Children comprehended the modeled words better in the follow-in condition,  $t(9) = 2.41$ ,  $p < .05$ . In this condition, the children averaged correct responses on 50% of the trials overall: the mean score was 4.0 out of eight trials per child per condition (two words for four sessions). This proportion was significantly above the chance performance of 25% (assuming children always picked an object, which they did not),  $p < .05$ , whereas the proportion of correct responses in the direct condition was not. There were no statistically reliable effects found for either of the production measures, which were both quite low in both conditions; of the 40 trained words (four per child, 10 children) there were only 14 productions, and these came from only five children.

Due to illnesses and absences, only seven of the 10 children could be given follow-up testing within a few days of the 2-week interval. Again, children spoke very little, and so neither production measure produced differences. (However, it should be noted that all of the three productions recorded in the follow-up session were from the follow-in condition.) Analysis of the comprehension task produced a significant difference in favor of the follow-in condition: 64% to 36%,  $t(6) = 4.58$   $p < .05$ . Analysis of individual subjects confirmed this trend: six of the seven children had better comprehension scores in the follow-in condition, and the other child had equal scores in both conditions. The probability of this occurring by chance alone is less than .05, sign test.

### Discussion

The main finding of the training study was that the follow-in strategy produced greater word learning, as measured by comprehension, than the direct condition. The very small amount of production does not permit firm conclusions. Though it is possible that more training would have produced more productions, other lexical training studies have obtained results with this amount of training (cf. Schwartz & Terrell, 1983). More likely, the small amount of production was probably due to the children's general shyness alone with strangers.

The results of this study help to explain those of Study 1. By themselves, the correlations of that study could be explained if it were the case that the child was (a) particularly attracted to some objects, (b) thus "primed" to learn their names, and (c) played with these objects most often. If this were the case it would mean that when mothers followed in, it would most often be attractive objects the child was playing with, and thus conditions for word learning would be maximal. Objects named in a directive manner would be those of little or no interest (the child was not playing with them) and so learning conditions would be less favorable. The results of Study 2, however, cannot be explained in this way since in this study objects were randomly assigned to conditions. Together, then, the results of the two studies are most economically explained by positing a facilitative effect of joint attentional processes.

Another possible explanation for the correlations of Study 1 is Nelson's (1981) functional hypothesis. It is possible that children with more directive mothers were learning that the primary function of language is social-regulative, and thus they were less interested in learning object names. Children of mothers who more often followed-in to their attention attributed to language more cognitive sig-



nificance and thus learned more object labels. This interpretation is not plausible in Study 2, however. In this study the same child learned words differentially depending on how they were presented. This could not be the product of one overall hypothesis about the functional significance of language. It is of course possible that Nelson's hypothesized mechanism is at work in the real world (and Study 1), while the current finding is a laboratory phenomenon; or it is possible that the two mechanisms are both operative, though on different levels. Once again, however, the most economical explanation of the two studies together is in terms of joint attentional processes.

### General Discussion

The current studies, in combination with findings of previous research, suggest that joint attention is important to early language acquisition in two ways. First, relatively extended episodes of joint attentional focus between adult and child provide important non-linguistic scaffolding for the young child's early linguistic interactions. This effect seems to extend well into the second half of the child's second year of life, beyond the very earliest stages of communicative development where most previous research has concentrated (e.g., that of Bruner and his colleagues). Further, what happens in these episodes seems to be of special importance for acquiring new language. Keith Nelson (e.g., 1982) has argued that, in general, when learning conditions are favorable the child's acquisition of novel linguistic structures may often be based on a single, or at most a very few, adult exemplars of that structure. The results of the current study suggest that for the initial phases of lexical development, relatively extended episodes of joint attentional focus between child and adult may constitute an important part of such conditions. This is presumably because such episodes are periods when the child is attentive, motivated, and best able to determine the meaning of her mother's language (cf. Ninio & Bruner, 1978; Ratner & Bruner, 1978).

Within joint attentional episodes, it would seem to be important that the adult talk about the object on which the child is focused, rather than constantly trying to redirect the child's attention. Roth (1985) has shown, in fact, that when mothers follow into their child's attentional focus, they are more likely to elaborate semantically on previous child utterances. In the current interpretation, the important factor in all cases is the relative

ease with which the child is able to establish the attentional focus of the adult and thus the referential context of her language. It is interesting to note in this regard that something very similar to this also operates at the level of conversational interaction. Olsen-Fulero (1982) has demonstrated that directiveness has an adverse effect on early mother-child conversations. If conversational topic may be thought of as analogous to an object of joint visual attention, then these results parallel those of the current study.

One final point should be made. All of the measures in the current study were of visual, not auditory, attention. This is quite simply because visual attention is most easily observable. It is possible, for example, that the child was indeed attending aurally to an object when she was coded as not attending—for example, when the mother shook a rattle that the child recognized. It is also possible that the child was not attending aurally to the mother's language in some cases—a situation not dealt with in the current study. Undoubtedly, a systematic account of auditory attention is necessary for a thorough understanding of the role of attentional factors in the language acquisition process.

Individual differences in early language acquisition present a challenge and an opportunity for researchers. As Katherine Nelson (1981) has pointed out, explaining these differences may play a crucial role in discovering the basic cognitive and social process that underlie language development. Thus far explanations have centered on such factors as cognitive style differences among children (Bretherton, McNew, Snyder, & Bates, 1983), social-interactional differences among mother-child dyads (Nelson, 1973), and differences among the social environments of children (i.e., the amount of interaction with fathers, siblings, peers, strangers, etc.; see Mandle & Tomasello, in press). It is safe to assume that each of these has some role to play. In this study we have attempted to identify and explore another set of factors that, like the others, may be fundamental for language acquisition and at the same time contribute to individual differences. Joint attentional processes are clearly worthy of future research attention.

### References

- Bakeman, R., & Adamson, L. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*, 1278–1289.
- Bates, E. (1979). *The emergence of symbols: Cogni-*

- tion and communication in infancy. New York: Academic Press.
- Bretherton, I., McNew, S., Snyder, L., & Bates, E. (1983). Individual differences at 20 months: Analytic and holistic strategies in language acquisition. *Journal of Child Language*, *10*, 293–320.
- Bruner, J. (1981). The pragmatics of acquisition. In W. Deutsch (Ed.), *The child's construction of language* (pp. 35–56). New York: Academic Press.
- Bruner, J. (1983). The acquisition of pragmatic commitments. In R. Golinkoff (Ed.), *The transition from prelinguistic to linguistic communication* (pp. 27–42). Hillsdale, NJ: Erlbaum.
- Bruner, J. (1985). *Child's talk: Learning to use language*. New York: Norton.
- Della Corte, M., Benedict, H., & Klein, D. (1983). The relationship of pragmatic dimensions of mothers' speech to the referential-expressive distinction. *Journal of Child Language*, *10*, 35–44.
- Mannle, S., & Tomasello, M. (in press). Fathers, siblings, and the bridge hypothesis. To appear in K. E. Nelson & A. van Kleeck (Eds.), *Children's language, Volume 6*. Hillsdale, NJ: Erlbaum.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, *38*(1–2, Serial No. 149).
- Nelson, K. (1981). Individual differences in language development: Implications for development and language. *Developmental Psychology*, *17*, 170–187.
- Nelson, K. E. (1982). Theories of the child's acquisition of syntax. *Annals of the New York Academy of Sciences*, *345*, 45–69.
- Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language*, *5*, 1–16.
- Olsen-Fulero, L. (1982). Style and stability in mother conversational behavior. *Journal of Child Language*, *9*, 543–564.
- Ratner, N., & Bruner, J. (1978). Games, social exchange, and the acquisition of language. *Journal of Child Language*, *5*, 391–402.
- Roth, P. (1985). *Timing and function of maternal speech to 12-month-olds*. Manuscript submitted for publication.
- Schwartz, R., & Terrell, B. (1983). The role of input frequency in lexical acquisition. *Journal of Child Language*, *10*, 57–66.
- Tomasello, M., Mannle, S., & Kruger, A. (1986). The linguistic environment of one to two year old twins. *Development Psychology*, *22*, 169–176.
- Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*, 197–212.