Three is not always a crowd: contexts of joint attention and language

JOANN P. BENIGNO, LAURA CLARK and M. JEFFREY FARRAR

Journal of Child Language / Volume 34 / Issue 01 / February 2007, pp 175 - 187
DOI: 10.1017/S0305000906007732, Published online: 25 January 2007

Link to this article: http://journals.cambridge.org/abstract_S0305000906007732

How to cite this article:

Request Permissions : Click here
NOTE

Three is not always a crowd: contexts of joint attention and language*

JOANN P. BENIGNO
Institute of Child Development, University of Minnesota

LAURA CLARK AND M. JEFFREY FARRAR
University of Florida

(Received 12 September 2005. Revised 5 June 2006)

ABSTRACT

This study examined 32 children’s (M age = 1;8 years) engagement in joint attention (JA) and the relation between JA and vocabulary size across mother–child (MC) and mother–child–sibling (MCS) contexts. In the MCS context, mothers engaged in JA more with one child than both children; they engaged in less JA with target child than they did in the MC context. JA style was generally unrelated across the contexts. Coordinated JA and children’s vocabulary were significantly related only for the MCS context. Findings suggest the number of social partners influences JA dynamics and multi-child contexts can be positive language learning environments.

INTRODUCTION

Between 9 and 12 months of age, important milestones emerge in children’s early social cognitive development such as gaze following, social referencing, joint engagement, imitative learning, and the use of gestures to direct...

[*] Analyses and the preparation of this manuscript were supported by a National Institute of Mental Health National Research Service Award awarded to the Institute of Child Development at the University of Minnesota (T32 MH 15755). Portions of this work were completed in partial fulfillment of the requirements for a senior honors thesis by the second author at the University of Florida. Further correspondence regarding this paper may be addressed to Joann P. Benigno, PhD, Ohio University, School of Hearing, Speech, and Language Sciences, W218 Grover Center, Athens, OH 45701 or via e-mail at benigno@ohio.edu.

We would like to thank all of the families who participated in the study. We would also like to thank the members of the Child Language Lab for assisting with data collection, transcription, and coding, particularly Anisa Mohammed and Zachalis Rodriguez. Finally we thank Patrick J. Carroll and Scott Miller for their comments on earlier versions of this paper.
others’ attention (Tomasello, 1999). Central to these behaviors is the infant’s ability to coordinate their attention between another person and an object or event, known as joint attention (JA; Bakeman & Adamson, 1984). Assessments of infants’ JA typically occur within the context of dyadic social interactions with caregivers, such as their mothers. However, the majority of children’s social interactions occur in the presence of others, such as peers and older siblings. Although empirical precedence exists for the examination of JA in multi-child contexts (Tomasello, Mannle & Kruger, 1986; Barton & Tomasello, 1991), questions remain regarding how features of such arrangements relate to early language skills in multi-child contexts. The goals of the present study were to examine and compare: (1) the characteristics of JA episodes within mother–child–sibling (MCS) contexts and mother–child (MC) contexts; and (2) whether the relation between JA and reported vocabulary size differed across the two interactive contexts.


However, JA is not always necessary for language learning to occur. In fact, children’s language learning also occurs in more incidental contexts, such as learning from overheard speech (Akhtar, Jipson & Callanan, 2001). Within the context of family interactions, Oshima-Takane and colleagues’ (e.g. Oshima-Takane, Goodz & Derevensky, 1996; Oshima-Takane & Robbins, 2003) findings suggest that toddlers benefit from overhearing the conversations shared between their older siblings and their mothers. One specific linguistic benefit is reflected in young children’s use of personal pronouns. Additional benefits from multi-party speech include peer talk, which is known to influence children’s pragmatic communicative abilities (Blum-Kulka & Snow, 2002). From a cross-cultural perspective, children in non-Western cultures are not primarily exposed to dyadic interactions from birth, yet they still acquire language (Lieven, 1994; Rogoff, 2003). Even in Western cultures, many children attend daycare and as a result, the majority of their language learning opportunities are likely occurring in multi-child contexts (NICHD Early Child Care Research Network, 2000).

Although children are capable of learning language through less explicit multi-partner contexts, some evidence suggests that multi-child contexts negatively impact language learning because the caregiver’s attention is split
amongst the children participating in the interaction (Jones & Adamson, 1987; Tomasello et al., 1986; Barton & Tomasello, 1991). Tomasello et al. (1986) found that twins engaged in dyadic JA interactions with their mothers far less frequently and for shorter lengths of time than singletons did. Relative to singletons, twin children demonstrated slower language development; they produced fewer words, had smaller vocabulary sizes, engaged in fewer conversations, and took fewer turns per conversation. Mothers who interacted with their twins also engaged in shorter conversations, imitated their child’s responses more often, and were less likely to follow into their child’s attentional focus. Further, infants produce fewer utterances when interacting with both their mother and sibling than when interacting with their mother alone (Jones & Adamson, 1987). Although mothers produced the same number of utterances across both contexts, the number of utterances directed toward the infants was significantly less in the triadic context.

Although some evidence suggests that additional social partners, such as siblings, negatively impact younger children’s ability to learn language, there is still reason to believe that multi-child contexts are positive language learning environments. Barton & Tomasello (1991) investigated the relations among age, JA, and conversational turns in mother–infant–sibling triads containing infants who were 1;7 versus 2;0 years. Infants as young as 1;7 years were able to participate in JA as well as conversational episodes with both their mothers and their siblings. However, those who were 2;0 years engaged in significantly more triadic JA and conversational episodes. The mean proportion of mother–sibling JA episodes exceeded mother–infant–sibling JA episodes only for triads containing an infant of 1;7 years. Regardless of the infant’s age, infant–sibling JA episodes were very infrequent. Conversations involving all three participants were almost three times longer and contained a greater amount of conversational turn-taking than either mother–infant or mother–sibling conversations.

The present study
In sum, the research reviewed (particularly Barton & Tomasello, 1991) suggests that triadic interactions have unique dynamics that distinguish them from dyadic interactions and that multi-child contexts provide a rich language learning environment. However, questions still remain regarding the quality and consequences of JA within MC and MCS contexts. Since no studies directly compared JA in MC and MCS contexts, we do not know which context promotes more frequent and longer-lasting JA episodes between children and their mothers. Within the MCS context, we predicted that: (1) mothers would spend more time engaged in JA dyadic interactions (with one of their children) than engaged in triadic interactions (with both of
their children); (2) mother–sibling JA episodes would be the most frequent and longest lasting type of JA episode and child–sibling JA episodes would be the least frequent and shortest type of JA episode; and (3) when all three members engaged in collective JA, the frequency and duration of coordinated JA would be significantly greater than the frequency and duration of passive JA.

The second aim of the study was to compare the mother–child–sibling MCS context to the mother–child MC context and determine whether similarities exist in incidence of coordinated and passive JA within and between the contexts. Across contexts, we predicted that: (1) JA episodes occurring in the MCS context would be more frequent and longer in duration than those in the MC context; and (2) families would use similar types of JA engagement across both contexts.

The final aim of the study was to examine the relation between JA and vocabulary size across the two interactive contexts for the young target child. The relative contribution of JA in these contexts to language development is unclear for several reasons. Studies investigating language-JA links in MCS and MC contexts vary in the measures used to index language. Those studying MCS contexts primarily used conversational measures (e.g. conversational turns), whereas those examining MC contexts have primarily focused on vocabulary size. The current investigation addresses these issues by directly comparing families interacting in MC contexts to MCS contexts within the same family. While we anticipated JA-vocabulary links in both contexts, we did not have a prediction about the relative strength of these relations. Although strong support exists for the relation between JA and vocabulary size in MC contexts, we also believed that the MCS context could foster a stronger positive relation between JA and vocabulary size because it is reflective of children’s typical learning environment of growing up in a multi-child context.

**METHOD**

**Participants**

Thirty-three families each consisting of the target child, an older sibling (preschooler), and a mother participated in this study. These families were primarily Caucasian and from middle- to upper-middle-class backgrounds. Eighty-eight percent of the mothers (\( M_{\text{age}} = 34.5; \ S.D. = 3.0 \) years) held a bachelor’s degree or higher and 61% were stay-at-home mothers. Sixteen of the older siblings were girls and 17 were boys (\( M_{\text{age}} = 4.7; \ S.D. = 0.5 \) months), and 26 of the target children were girls and 7 were boys (\( M_{\text{age}} = 1.8; \ S.D. = 0.4 \) months). Three of the sibling pairs were brother–brother, 12 were sister–sister, 14 were older brother–younger sister, and 4 were older sister–younger brother.
Materials

Free-play interactions. Two video cameras mounted on tripods were used to videotape the free-play interactions. A 50” by 60” blanket was placed on the ground to indicate the range of the video cameras. One of three sets of toys was used for each interaction. Each set of toys included the following types of items previously used by Carpenter et al. (1998): bucket, blocks, ball, doll or animal, stacking cups, small figurines, toys with wheels, picture books, and small rolling objects. The sets of toys were counterbalanced across all families.

Language. Mothers completed the MacArthur-Bates Communicative Development Inventories (MCDI; Fenson et al., 1993), which served as an index of the younger child’s communicative abilities. Two forms of the assessment were used due to the infants’ range of ages from 1;2 to 2;4. Mothers of children younger than 1;4 years (n=4) completed the MacArthur Inventory for Communicative Development: Words and Gestures and mothers of children older than 1;4 (n=27) completed the MacArthur Inventory for Communicative Development: Words and Sentences.

Procedure

The current study was part of a larger study of young children’s emerging social cognitive development in family interactions. Each family participated in two separate sessions in their homes. Sessions occurred 1–2 weeks apart. As part of the first session, mothers engaged in a 10-minute free-play interaction with the target child. Older siblings completed tasks relevant to other portions of the study during the filming of the MC interaction phase. As part of the second session, mothers engaged in a 10-minute free play with the target child and his or her preschool-aged older sibling. During the free-play interactions, a researcher placed one of the three sets of toys on a blanket and instructed the participants to remain on the blanket while playing. Research assistants were present throughout the 10-minute interactions for filming purposes.

Coding of JA. The coding procedures of Bakeman & Adamson (1984) were used to classify episodes of coordinated JA and passive JA across both contexts. Coordinated JA was defined as active engagement with the same object shared between all communicative partners involved, with awareness of the presence of the other partner(s) indicated by eye contact/alternating eye gaze, verbal response, or some other overt behavior. The interaction had to last for a minimum of three seconds. Interactions were defined as passive JA when two communicative partners were both actively involved with the same object for at least three seconds, but there was no indication of awareness of the presence of the other partner(s). A JA episode ended
when one of the partners changed his or her attentional focus for three or more seconds and did not return his or her attention to the original object within 10 seconds. Further within the MCS context, different types of coordinated and passive JA could occur for different combinations of social partners. Specifically, within the MCS context, both types of JA could occur for mother–child, mother–sibling, and child–sibling dyads, as well as mother–child–sibling triads.

**Reliability of JA coding.** Fifteen percent of the videotaped free-play sessions were independently coded by a second coder for JA episodes. The overall agreement of the frequency and duration of JA episodes was 91% and 89% respectively. Cohen’s kappa for agreement on overall JA episodes (collapsed across coordinated and passive JA) was 0.80. Cohen’s kappa for the agreement of coordinated versus passive JA episodes was 0.77. According to Bakeman and Gottman (1986) Cohen’s kappas in this range are considered a ‘good to excellent’ level of agreement.

**RESULTS**

The sample used for data analysis included 32 families since one family did not complete the videotaped interaction procedure.

**Joint attention in the MCS context**

The first aim of the study was to examine the characteristics of JA episodes occurring in the MCS context. Table 1 presents the means and standard deviations for the frequencies, durations, and mean durations (duration/frequency) of each JA type. Child–sibling passive and coordinated JA episodes and mother–sibling passive JA episodes in the MCS context were not included in the analyses due to the very limited number of occurrences of these events.

Overall, the frequencies and time durations of passive and coordinated JA episodes did not differ within the MCS. However, as predicted, mother–child coordinated JA episodes lasted approximately three times longer than mother–child passive JA episodes. Analyses revealed significant differences in the frequencies, duration of time, and mean duration of time that mother–child pairs within the MCS spent in coordinated JA versus passive JA, all $t$s > 2.93, all $p$s < 0.01.

As predicted, mothers spent over four times longer engaged in dyadic JA episodes with one child than in triadic JA episodes with both of their children together, $t(31) = 8.41, p < 0.001$ ($M = 393.14, s.d. = 139; M = 80.83; s.d. = 14.29$, respectively). Although mothers engaged in coordinated JA with the siblings a significantly greater number of times than with the target child, $t(31) = 2.36, p < 0.05$, the durations of these episodes were not
significantly different. Lastly, the mean duration of MC coordinated JA episodes was longer than MS coordinated JA, but this difference failed to reach statistical significance, \( t(31) = 1.93, p = 0.06 \).

**Joint attention: similarities and differences between MCS and MC Contexts**

The second aim of the study was to directly compare JA episodes in the MCS context to the MC context. Table 2 presents the means and standard deviations for the frequency, duration, and mean duration JA measures for both contexts. Contrary to our predictions, the total frequency of JA involving the target child was significantly greater in the MCS context than in the MC context, \( t(31) = 2.43, p < 0.05 \). No significant differences emerged in the frequency of mother–child passive or coordinated JA between the two interactive contexts. However, in line with our predictions, the total duration of JA as well as the mean duration per JA episode involving the target
child was significantly greater in the MC context, both $t_s > 4.53$, $p_s < 0.001$. Mother–child dyads in the MC context spent well over three times longer engaged in passive JA and over 1.5 times longer engaged in coordinated JA relative to their engagement in JA in the MCS context. Differences emerged between the two contexts in the total durations as well as the mean

**TABLE 2. Mean frequency and duration of JA in mother–child–sibling triads and mother–child dyads**

<table>
<thead>
<tr>
<th>Joint attention (JA) type</th>
<th>Triads ($n=32$)</th>
<th>Dyads ($n=32$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of mother–child passive JA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>2.72$^a$</td>
<td>4.41$^a$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(2.74)</td>
<td>(4.04)</td>
</tr>
<tr>
<td>Frequency of mother–child coordinated JA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>4.97$^a$</td>
<td>3.91$^a$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(2.60)</td>
<td>(3.18)</td>
</tr>
<tr>
<td>Total frequency of JA with child</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>10.72$^a$</td>
<td>8.31$^b$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(3.55)</td>
<td>(3.80)</td>
</tr>
<tr>
<td>Duration of mother–child passive JA (in seconds)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>53.75$^a$</td>
<td>179.88$^b$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(46.00)</td>
<td>(167.59)</td>
</tr>
<tr>
<td>Duration of mother–child coordinated JA (in seconds)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>165.53$^a$</td>
<td>265.56$^b$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(110.41)</td>
<td>(207.24)</td>
</tr>
<tr>
<td>Total duration of JA with child (in seconds)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>313.53$^a$</td>
<td>445.44$^b$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(94.72)</td>
<td>(127.01)</td>
</tr>
<tr>
<td>Mean duration of mother–child passive JA (in seconds)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>14.07$^a$</td>
<td>41.18$^b$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(16.10)</td>
<td>(53.32)</td>
</tr>
<tr>
<td>Mean duration of mother–child coordinated JA (in seconds)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>27.26$^a$</td>
<td>69.51$^b$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(33.24)</td>
<td>(59.00)</td>
</tr>
<tr>
<td>Mean duration of total JA mother–child episodes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>29.25$^a$</td>
<td>69.37$^b$</td>
</tr>
<tr>
<td>$(s.d.)$</td>
<td>(8.83)</td>
<td>(52.14)</td>
</tr>
</tbody>
</table>

*Note:* For each JA type, means within rows with different superscripts differ at $p < 0.05$. 
durations of the specific JA types (passive and coordinated), all $t_s > 2.50$, all $p_s < 0.05$.

To determine whether families adopted similar styles of JA engagement across the interactive contexts, we examined the relations among all JA measures between the two contexts. One-tailed tests were used, as our hypotheses were all directional. The only significant relations that emerged between the MC and MCS contexts were the frequency of total JA episodes involving the target child, $r = 0.52$, $p < 0.01$ and the mean duration of total JA episodes involving the child, $r = 0.31$, $p < 0.05$. Once the child’s age had been controlled, the relation between the frequency of total JA remained significant, $r = 0.49$, $p < 0.01$, but the relation between total mean duration of JA across contexts was no longer significant ($r = 0.17$). Contrary to our predictions, no significant relations were found between contexts for the overall duration and mean duration measures, nor for the more specific forms of JA.

**Joint attention and language**

The final aim of the study was to investigate relations between JA and vocabulary size within the MC and MCS contexts. One family did not complete the MacArthur Communicative Development Inventory; thus, 31 families were included in the language analyses. Percentile rank scores were used since two different versions of the MCDI were employed in the current sample. Mean percentile rank in the current sample was 51.32 ($s.d. = 30.11$). Again, one-tailed tests were used, as our hypotheses were all directional.

In the MCS context, as predicted, the durations of mother–child coordinated JA episodes and total JA episodes significantly related to the target child’s standard score for reported vocabulary size, $r = 0.46$, $p < 0.01$ and $r = 0.30$, $p < 0.05$ respectively. The pattern of findings was similar when examining the relations between vocabulary size and the mean duration of mother–child coordinated JA episodes and total JA episodes involving the toddler, $r = 0.29$, $p = 0.05$ and $r = 0.30$, $p = 0.05$. In contrast to our predictions, no relation emerged between the total duration of coordinated JA and vocabulary size within the MC ($r = 0.08$). Further, no significant relations between all passive JA measures and reported vocabulary size for both the MC and MCS contexts (all $p_s > 0.09$) were found.

To further explore this relation, a simultaneous regression analysis was used to determine whether MC coordinated JA within the MCS context uniquely predicted children’s percentile scores on the MCDI. The following variables were entered: duration of mother–child coordinated JA in the MCS context, duration of mother–child coordinated JA in the MC context, and total duration of coordinated JA involving the target child in the MCS context.
context. When the children’s percentile scores on the MCDI were regressed onto these three measures, only the duration of mother–child coordinated JA in the MCS context was a significant predictor of vocabulary size, $\beta = 0.42$, $t = 2.58$, $p = 0.01$. In contrast to previous studies, this finding suggests that MCS contexts are beneficial for vocabulary development.

**DISCUSSION**

Our examination of MCS and MC interactions supported several of our key hypotheses regarding the types of JA that emerged in both contexts as well as the relation between the target children’s JA and reported vocabulary size. The overarching message emerging from the present investigation is that multi-child contexts can be positive learning environments in fostering early social-cognitive and language development.

As hypothesized, mothers spent significantly more time engaged in JA with one of their children than with both of their children. This finding resonates with that of Barton & Tomasello (1991), who found that within MCS contexts, the average length of time of MS JA episodes was greater than the average length of time of MCS JA episodes. One possible explanation for why mothers spent more time engaged in JA with one child than with both children might relate to the increased social demands that characterize triadic contexts (Benigno & Ellis, 2004). When interacting with both children, mothers might have been more concerned with satisfying the particular wants of each child than trying to direct both children to engage in the same activity.

In the MCS context, mothers and their toddlers spent significantly more time engaged in coordinated JA than in passive JA. This was expected since by 1;8 years of age (the mean age of children in our study), not only are the children’s JA skills more refined, but they also have emerging language structures to support and facilitate JA with their social partners. However, mothers engaged in coordinated JA with the target child significantly less often than with the sibling. Thus, despite the target child’s emerging language, older siblings were still more adept at engaging their mothers’ attention not only through sharing attention, but by also participating in more complex JA interactions involving conversation. Also in concert with previous investigations (Bakeman & Adamson, 1984; Barton & Tomasello, 1991), very few child–sibling JA episodes occurred within the MCS context.

As hypothesized, when interacting with the target child in the MC context, mothers spent more time engaged in all comparable types of JA than when they interacted with the child in the MCS context. This was expected since mothers had to share their time with both children during MCS interactions, in contrast to the MC interactions, where
their attention was undivided. Interestingly, however, for coordinated JA, the most sophisticated form of JA, no differences emerged between the contexts, perhaps suggesting that these target children were equally capable of engaging in this form of JA across contexts. Relative to prior investigations examining interactions within MCS contexts, our study examines the nature and quality of JA both within and across MCS and MC contexts.

Although the total frequency as well as the mean duration of JA types (both passive and coordinated combined) involving the target child significantly related across the MC and the MCS context, contrary to our hypothesis, we found that families’ styles of JA engagement were for the most part dissimilar across contexts. The findings for the total frequency of JA suggest that the active engagement of the infant in JA within the MC context extends to active engagement in JA within the larger MCS context. An alternative interpretation would suggest that since no other relations emerged between the contexts for the duration of JA as well as any other types of JA, differences in the dynamics of the contexts may prevent families from using the same styles of interaction in both contexts. Perhaps some aspects of JA are contextually based, since mothers may have different goals in each context. For instance, it may be the case that mothers’ goals in the dyad are oriented towards teaching opportunities, relative to the triadic contexts, where the goals are to manage the flow of the interaction (Benigno & Ellis, 2004).

A particularly important result was the finding that only JA in the MCS context uniquely predicted vocabulary size. The finding that dyadic JA was not related to infants’ language abilities differs from previous research indicating that children who participate in more JA interactions with their mothers have larger vocabularies (Tomasello & Todd, 1983; Tomasello & Farrar, 1986) and can comprehend and produce more language (Carpenter et al., 1998). Specifically, Tomasello and Todd (1983) found that mother–child dyads who engaged in longer bouts of JA and spent a higher proportion of time in equal lead bouts of JA contained children with larger vocabularies.

Our findings also qualify previous research which suggests that the presence of an additional child negatively impacts children’s emerging vocabulary (Tomasello et al., 1986; Woollett, 1986; Jones & Adamson, 1987). However, our findings extend previous research which suggests that MCS contexts can have beneficial effects on infants’ conversation style (Dunn & Shatz, 1989; Barton & Tomasello, 1991). The findings are surprising, particularly since the benefits revealed by previous studies all involved pragmatic skills, not strict linguistic skills, such as vocabulary size.

The relations emerging between coordinated JA and vocabulary size cannot be explained simply by the amount of time spent engaged in JA,
because the overall amount of JA, particularly the frequency and duration of coordinated JA shared between mothers and toddlers, was significantly greater in the dyadic than the triadic context, not the other way around. Perhaps JA in MCS contexts, and not dyadic contexts, is more strongly related to the child’s vocabulary size because triadic contexts may be a more reasonable reflection of what these children experience in their multi-child family structures. Thus, those children capable of engaging in JA in these more challenging MCS contexts (which probably reflect their everyday experiences) likely possess the social-cognitive abilities that would facilitate their vocabulary acquisition (Baldwin, 1995). Of course, in order to determine the causal nature of this relation, longitudinal investigations are warranted.

In conclusion, multi-child contexts are important language learning environments for many children around the world. The results of the current study revealed that participation in JA episodes within MCS contexts has positive effects on emerging language skills. Our findings further imply that there are features unique to MCS contexts which lead to variations in the styles of JA employed across contexts. These variations suggest that the number of social partners influences the expression of young children’s social cognition and as well the relation between their social cognition and their emerging language abilities.

REFERENCES


186


