Joint Engagement and the Emergence of Language in Children with Autism and Down Syndrome

Lauren B. Adamson · Roger Bakeman · Deborah F. Deckner · MaryAnn Romski

Abstract Systematic longitudinal observations were made as typically developing toddlers and young children with autism and with Down syndrome interacted with their caregivers in order to document how joint engagement developed over a year-long period and how variations in joint engagement experiences predicted language outcome. Children with autism displayed a persistent deficit in coordinated joint attention; children with Down syndrome were significantly less able to infuse symbols into joint engagement. For all groups, variations in amount of symbol-infused supported joint engagement, a state in which the child attended to a shared object and to language but not actively to the partner, contributed to differences in expressive and receptive language outcome, over and above initial language capacity.

Keywords Autism · Down syndrome · Joint attention · Language · Parent–child interaction

The early development of joint engagement and the acquisition of language are usually intertwined. As infants acquire joint attention skills, they gain entrance into “primordial sharing situations” (Werner and Kaplan 1963) that serve as “zones of proximal development” (Vygotsky 1978) where caregivers can facilitate their introduction to symbols during affect-laden and intention-filled social interactions (Adamson 1996; Hobson 2000; Tomasello 1995). Then, as the child acquires a vocabulary, the scope of joint engagement increasingly expands as the focus of shared attention is displaced from present objects to symbols that refer to them, to future and past events, and to internal states (Adamson and Bakeman 2006).

Developmental disorders may disrupt the mutual relation between joint attention and language. Problems sustaining joint attention may alter and even curtail a toddler’s access to language-facilitating interactions. For example, the early deficits in joint attention skills in young children with autism that have been documented in home videotapes (e.g., Baranek 1999; Werner and Dawson 2005), standardized tests (Lord 1995; Mundy et al. 1990; Wetherby et al. 2007), screening questionnaires (Baron-Cohen et al. 1996; Robins et al. 2001), and laboratory-based studies (McArthur and Adamson 1996; Sigman and Ruskin 1999; Stone et al. 1997; see Leekam and Moore 2001, for a review) likely play a pivotal role in their problems acquiring and using language (Charman 2004; Dawson et al. 2004; Happé 1998; McDuffie et al. 2005; Sigman and Ruskin 1999; Smith et al. 2007; Toth et al. 2006). Thus, although older estimates that almost 50% of children with autism remain nonverbal (Tager-Flusberg 1994) are being revised downward (Lord et al. 2004), difficulties acquiring and using language remain one of the most daunting challenges facing young children with autism (Rice et al. 2005).

In addition, problems acquiring language may limit the scope of joint engagement. As a group, young children with Down syndrome may be especially prone to this limitation because they so often have more difficulty acquiring
expressive language than their level of cognitive delay might suggest (Chapman 1997, 2003; Rice et al. 2005; Roberts et al. 2007; Sigman and Ruskin 1999). Thus, even though they acquire joint attention skills (Sigman and Ruskin 1999), produce referential points (Franco and Wishart 1995) and coordinated joint attention looks (Kasari et al. 1995), and slowly become able to engage in coordinated joint attention during social interactions (Legerstee and Weintraub 1997), a relative lack of language facility may confine the topic of social interaction to immediate, concrete events.

In the current study, we sought to disentangle the effect of problems related to joint attention development from those arising from language delay by documenting the joint engagement experiences of young children with autism and children with Down syndrome and discerning what aspects of these experiences best predict language outcome. This research builds upon our longitudinal study of the emergence of symbol-infused joint engagement in typically developing toddlers (Adamson et al. 2004). Our basic strategy was to systematically describe engagement states of two new samples, one of young children with autism and one of young children with Down syndrome, who we observed repeatedly as they interacted with their mothers in a range of communicative contexts. This design allows us both to gather systematic observations about how each disorder affects caregiver-child communication and, through a comparative analysis, to use our findings to reflect on broader issues of how symbol-infused joint engagement emerges when joint engagement or language skills crucial to this state are compromised.

Our conceptualization of shared engagement (Adamson and Bakeman 1991; Adamson and Chance 1998) led us to formulate a coding scheme that captured two key distinctions in a child’s experience during periods of joint engagement. The first distinction is between supported and coordinated joint engagement and focuses on whether or not the child is explicitly attending not only to a shared topic but also the partner. Coordinated joint engagement typically gradually emerges between 9- and 15-months of age, so that by 18-months toddlers can sustain periods during which they punctuate actions on a common topic with explicit communicative actions such as well-timed glances at their partner’s face (Bakeman and Adamson 1984; Legerstee and Weintraub 1997). However, as the images of “zone of proximal development” evoke, a child may also actively share events and objects without explicitly acknowledging the social partner. Such periods of supported joint engagement occur often in mother-infant interactions (but not infant-infant interactions) by the middle of the first year (Bakeman and Adamson 1984), and they remain prevalent as language emerges between 18- and 30-months of age (Adamson et al. 2004). Over time the composition of this state may change as both the children (Bakeman and Adamson 1986) and the mothers (Adamson and Bakeman 1984) alter their actions, but maternal scaffolding of the interpersonal channel remains a constant. The second distinction characterizes the child’s attention to symbols. Although adult partners often speak during even their first encounter with a newborn (Rheingold and Adams 1980), infants typically begin to attend to and understand symbolic content and to produce symbolic acts late in their first year and then with increasing regularity and variety during their second year (Adamson 1996). Thus, initial periods of joint engagement are non-symbol-infused and then, as symbolic skills emerge, they become symbol-infused (Adamson et al. 2004).

These two distinctions generate four forms of joint engagement. The contrast between these forms can be illustrated by imaging four variations on a play sequence during which a child engages in an activity, putting pieces into a puzzle. In all four variations, the child permits the caregiver to also join in the activity (hence this is joint engagement, not solitary object engagement). An episode that would be coded as non-symbol-infused supported joint engagement might be described as follows: the child’s attention is attracted to the puzzle by the caregiver who places it in front of the child and hands the child a piece; the child readily accepts the piece and tries to place it in a puzzle as the mother assists by orienting the puzzle frame, naming the piece and verbally encouraging the child (“That’s a cow. You can put it over in the barn”) and pointing to the correct location; the child successfully places the piece and reaches for another one. Note that the child’s experience with the objects was enhanced by the mother’s scaffolding contributions even though the child did not explicitly communicate with her. For this sequence to be considered an episode of coordinated joint engagement, the child would also have to display explicit attention to the mother, for example, looking at her face when he accepts the puzzle piece, glancing towards her with a smile when the piece is correctly placed, or handing her the second puzzle piece and saying “your turn.” For the sequence to be considered symbol-infused, the child would have to act in a way that made it clear that he was attending to symbols as well as present objects and actions by, for example, speaking (“A horse!”) or following the mother’s specific statements (e.g., reorienting a piece when the mother says “turn it around”).

Our study’s first aim was to determine if and how autism and Down syndrome affected joint engagement during caregiver-child interactions. To this end, we compared the distribution of forms of joint engagement in 30-month-old children with autism and with Down syndrome to two points along the course of joint engagement development in typically developing toddlers: at 18 months when they have consolidated coordinated joint engagement and are just beginning to infuse symbols into joint engagement, and...
at 30 months when they are able to sustain periods of symbol-infused joint engagement. In light of prior studies, we predicted that joint engagement would occur less often in the group of children with autism than in either the typically developing or Down syndrome groups, regardless of their language status. We also anticipated that coordinated joint engagement would be especially problematic for children with autism but that supported joint engagement might be relatively spared and, when children with autism are verbal, they might infuse symbols into periods of joint engagement.

In contrast, we expected that, compared to typically developing toddlers of equivalent language facility, children with Down syndrome would have little problem sustaining joint engagement overall and coordinated joint engagement in particular. Indeed, we expected they might display relatively low levels of solitary object engagement (Cielinski et al. 1995; Legerstee and Weintraub 1997, but see Harris et al. 1996) and that the amount of supported joint engagement might be elevated because their partners might be more apt to scaffold object-focused attention (Cielinski et al. 1995; Harris et al. 1996; Roach et al. 1998). Still, we anticipated that their expressive language deficits might result in very low levels of symbol-infused joint engagement.

We also anticipated that there might be differences across groups that reflected diagnosis-specific difficulties with specific communicative functions. Although infants typically begin to engage in commenting by the end of their first year, young children with autism have been repeatedly shown to find this communicative function especially difficult (Adamson et al. 2001; Mundy 1995; Sigman and Ruskin 1999; Tager-Flusberg 1994; Wetherby and Prutting 1984). In contrast, compared to typically developing toddlers, young children with Down syndrome seem to find requesting objects or assistance with objects particularly difficult (Mundy et al. 1995). Thus we hypothesized that children with autism would have the most difficulty sustaining joint engagement in contexts that encouraged commenting rather than interacting or requesting and that children with Down syndrome might be less likely to sustain joint engagement in contexts that afforded requesting.

Cross-sectional comparisons can highlight differences between groups at crucial developmental transition points, but they do not reveal the trajectory of deviations from a typical milestone. Therefore, our second aim was to chart trajectories to test the hypotheses that a deficit in coordinated joint engagement and a relative sparing of supported joint engagement persist in autism, even if the child becomes increasingly verbal, and that a deficit in symbol-infusion might persist in young children with Down syndrome without an accompanying deficit in coordinated joint engagement. To this end, we tracked the distribution of forms of joint engagement over a year-long period. Drawing from a case study analysis of the paths taken by young children with autism (Adamson et al. 2005), we assigned our participants to groups based on language skill at the beginning and end of the year of observation so that we could compare developmental courses taken by children who were already verbal when we began to observe them with those who remained nonverbal and those who became verbal across the span of the study. Our primary hypotheses were first, that the amount of coordinated joint engagement would remain stable in all groups, and second, that children who became verbal shifted from resembling nonverbal to resembling verbal children in how often symbols infused their joint engagement. Thus, we expected that there would be some children with autism who would rarely sustain coordinated joint engagement but who would become verbal and infuse symbols into periods of joint engagement over the course of our year of observations and some children with Down syndrome who would often sustain coordinated joint engagement but would not become verbal and infuse symbols into joint engagement. These findings would suggest that sustaining periods of coordinated joint engagement may not be a necessary or sufficient step in the emergence of symbol-infused joint engagement.

Our final aim was to determine whether variations in joint engagement within the autism and Down syndrome samples would help predict language facility. In our study of typically developing toddlers (Adamson et al. 2004), we found that variations in amount of symbol-infused supported joint engagement, and not the amount of other forms of joint engagement, accounted for a significant amount of the variation in expressive and receptive language outcome at 30 months of age over and above their initial language level. This finding—that the facilitative effect of symbol-infused supported joint engagement is an especially potent context for early vocabulary acquisition—is consistent with views of language acquisition (e.g., Bruner 1983; Nelson 1996) that locate crucial learning in interactions that focus the child’s attention on shared objects and symbols. Moreover, it is in line with Bloom and Tinker’s (2001) provocative claim that toddlers who are less apt to engage with their caregivers in affective-filled interpersonal exchanges may acquire vocabularies more rapidly because they focus more often on the cognitively demanding process of relating symbols to shared objects. Thus, we hypothesized that variations in how often during a year-long period a child, regardless of diagnostic group, engages in symbol-infused supported joint engagement may contribute to differences in language abilities at the end of the observation period, even after controlling for the child’s initial language status. In contrast, we did not expect the amount of time spent in symbol-infused coordinated joint engagement to predict language outcome.
Method

Participants

The primary participants were 23 children with autism (20 boys) and 29 children with Down syndrome (19 boys) who were observed interacting with a parent (in all but two cases the mother) from two to five times over the course of a year. Their mean age at the first observation was 30.8 and 30.3 months (SD = 4.6 and 4.9) for autism (AU) and Down syndrome (DS) samples, respectively. Used for comparison were 56 typically developing (TD) toddlers who were observed between 18 and 30 months of age during a previous study (Adamson et al. 2004); their mean ages at the first and last observations were 18.1 and 30.0 (SD = 0.3 and 0.4), respectively. For the AU, DS, and TD samples, respectively: 83, 79, and 79% of the children were European American, 21 and 16% were African American, and 13, 0, and 4% were Hispanic; 65, 79, and 75% of the parents had earned at least a bachelor’s degree. English was the primary language in all homes. All of the children in the AU and DS samples except for one child with DS received clinical services such as speech therapy (96 and 97% of the AU and DS samples, respectively), occupational therapy (96 and 83%), and special needs preschool (83 and 62%) during at least some portion of the study period.

For the AU sample, 19 were observed five times, 1 three times, and 3 two times. Eligible participants were identified by three clinicians in our metropolitan area. When parents consented to participate, we administered the Autism Diagnostic Interview-Revised (ADI-R; Lord et al. 1994), an interview that assesses behavior in three core domains—social interaction, communication (verbal or nonverbal), and restricted or repetitive behaviors—and determines whether difficulties were evident before the child’s third birthday. The ADI-R results were consistent with the clinician referrals in all cases. Specifically, 22 children scored above cut-off on all three behavioral domains; 1 child scored above the cut-off on two of the scales and one point below cut-off on the nonverbal communication scale. For the DS sample, 15 were observed five times, 9 four times, 4 three times, and 1 two times. Recruitment was done using the network of referral sources for studies of young children with communication delays (Romsiki et al. 1997). For the TD sample, 54 were observed five times and 2 four times. During the first visit, all of the mothers reported that the child was full-term and healthy at birth and that they had experienced no health or developmental problems other than minor childhood illnesses during their first 18 months.

Recording Sessions

Participants were observed in a 4.6 × 3.1 m playroom in our laboratory during sessions that lasted approximately 2 h. Dyads were videotaped by two cameras placed behind one-way mirrored windows on opposite walls of the playroom as they engaged in semi-natural conditions using the Communication Play Protocol (CPP; Adamson et al. 2004). This arrangement allowed us to observe both the child and the parent as they moved around the playroom. The protocol contained six 5-min scenes that encouraged interactions in three communicative contexts: social interacting (sharing music; taking turns), requesting (assistance gaining toys from a high shelf, help playing with complicated toys), and commenting (sharing pictures, discussing objects in a container). The three communicative contexts lasted on average 620 s each (SD = 20), with no differences between groups. A manual for the CPP is available upon request.

Coding

As in Adamson et al. (2004), 11 engagement state codes that characterized the child’s active attention to people, objects, and symbols were defined. Of primary interest for this report were the four forms of joint engagement: non-symbol-infused supported, non-symbol-infused coordinated, symbol-infused supported, and symbol-infused coordinated (see Table 1). Two other non-symbol-infused states were included for comparison (unengaged and object). The remaining five codes (onlooking, person, symbol-only, person-symbol, and object-symbol) made the set of codes exhaustive, but in fact occurred relatively infrequently; consequently, for reliability we lumped them together as an other code and did not analyze them further. For subsequent analyses, consistent with our hypotheses, we computed the percentage of time each child devoted to symbol-infused joint engagement (pooling supported and coordinated symbol-infused); we also computed percentage of time devoted to supported joint engagement (symbol-infused or not), coordinated joint engagement (symbol-infused or not), solitary object play, unengaged, and other (the last five engagement state scores sum to 100%).

To establish interobserver reliability, at least 15% of the corpus for each sample was coded independently by two teams, masked as to which were reliability tapes. Pooled Cohen’s kappas (Cohen 1960), a statistic that assesses the reliability of a categorical scale while correcting for chance agreement (Bakeman and Gottman 1997), were .76, .69, and .71 for the AU, DS, and TD samples, respectively, values that Fleiss (1981) designated good to excellent.

Language Outcome

After the CPP at the fifth visit, we administered the Peabody Picture Vocabulary Test-III (PPVT-III; Dunn and Dunn 1997) and the Expressive Vocabulary Test (EVT; Williams 1997). Standardized measures of receptive and
expressive vocabulary skills, respectively. The PPVT was administered to 18, 22, and 53 and the EVT to 18, 22, and 54 AU, DS, and TD children, respectively. For the PPVT, 7, 10, and 1, and for the EVT, 5, 10, and 1, AU, DS, and TD children did not obtain a basal score and were assigned scores of 40, the lowest standard score for both tests; the lowest obtained score was 49 and 43, respectively.

Language Trajectory Groups

Trajectory groups consisted of children who remained relatively nonverbal, those who became verbal across the span of the study, and those who were already verbal when we began to observe them. One-hundred two children were classified: 20 AU, 28 DS, and 54 TD children (excluded were 3 AU children and 1 DS child with only two visits and 2 TD child with only one CDI score). Approximately 1 week before each visit, mothers filled out the MacArthur Communication Development Inventory (CDI; Fenson et al. 1993), which reports the number of words in the child’s expressive vocabulary. Trajectory groups were then based on the pattern of number of words reported (CDI scores) across visits. We classified children as minimally verbal if their CDI score never exceeded 50 at any visit, a size that has often been used as a marker for an acceleration or ‘spurt’ in word learning (e.g., Goldfield and Reznick 1990; Nazzi and Bertoncini 2003; 5 AU and 6 DS children met this criterion). We classified children as already verbal if their CDI score was 100 or higher at all visits (7 AU, 2 DS, and 8 TD children). We classified the remaining children as emerging verbal (8 AU, 20 DS, and 46 TD). Mean CDI scores for initial to final visits for emerging verbal children were 18–300, 34–219, and 49–478 for AU, DS, and TD children, respectively; comparable ranges for already verbal children were 286–554, 201–381, and 201–625.

Results

Effect of Diagnostic Group on Joint Engagement

Our first aim was to compare the joint engagement of children with autism, children with DS, and their typically developing peers. Thus, we computed time spent in symbol-infused joint engagement as well as time spent in supported joint engagement, coordinated joint engagement, solitary object play, unengaged, and other (the last five engagement state scores sum to 100%), for the AU and DS children’s 30-month visit and the TD children’s 18-month and 30-month visits. For these and subsequent analyses, AU30 and DS30 were compared with TD18 and not TD30 scores because the language level of TD children at 18 months more closely matched the language level of the other two groups at 30 months, and so provided a more appropriate comparison. The six engagement state scores were subjected to diagnostic (AU30, DS30, TD18) analyses of variance, N = 108 (TD30 means are provided for comparison). For each variable within diagnostic group, means that did not differ per Tukey post hoc test, p < .05, share a common subscript

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Engagement state codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement state</td>
<td>Definition</td>
</tr>
<tr>
<td>Unengaged</td>
<td>The child appears uninvolved with people, objects, events, or symbols</td>
</tr>
<tr>
<td>Object</td>
<td>The child is exclusively engaged with objects</td>
</tr>
<tr>
<td>Supported joint</td>
<td>The child and parent are actively involved with the same object or event, but the child is not actively acknowledging the parent’s participation</td>
</tr>
<tr>
<td>Coordinated joint</td>
<td>As is the case with supported joint engagement, the child and parent are actively involved with the same referent, and the child is actively and repeatedly acknowledging the parent’s participation, likely by visually referencing the parent at critical junctures in the interaction</td>
</tr>
<tr>
<td>Symbol-infused supported joint</td>
<td>The child and parent are engaged with the same referent and there is evidence that the child is actively attending to symbols, but the child is not explicitly attending to the parent; e.g., the parent may be assisting the child as she focuses on naming and manipulating pieces of a puzzle</td>
</tr>
<tr>
<td>Symbol-infused coordinated joint</td>
<td>The child is coordinating her attention between the parent and a shared referent, and the child is actively attending to symbols; the child may indicate her attention to the parent either by glancing toward her and/or by talking to her (e.g., Mommy, it’s your turn)</td>
</tr>
<tr>
<td>Other</td>
<td>Primarily Onlooking (child is watching the parent’s activity) and Person (child is interacting only with the parent) but also infrequent codes of Symbol-only, Person-symbol, and Object-symbol; see text for details</td>
</tr>
</tbody>
</table>

Table 2 Effect of diagnostic group on engagement state variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diagnostic group means</th>
<th>Diagnostic effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AU30</td>
<td>DS30</td>
</tr>
<tr>
<td>Symbol-infused</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Supported</td>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>Coordinated</td>
<td>5a</td>
<td>22b</td>
</tr>
<tr>
<td>Object</td>
<td>26a</td>
<td>12a</td>
</tr>
<tr>
<td>Unengaged</td>
<td>15b</td>
<td>7a</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Note: Means are percents. Effect sizes are from diagnostic group (AU30, DS30, TD18) analyses of variance, N = 108 (TD30 means are provided for comparison). For each variable within diagnostic group, means that did not differ per Tukey post hoc test, p < .05, share a common subscript.
between-subjects analyses of variance (see Table 2; TD30 means are included in Table 2 for comparison). Engagement state means are also presented in Fig. 1, separately for the three diagnostic groups, in a way that shows graphically how symbol-infused overlaps part of supported and coordinated joint engagement.

The results support our expectation that coordinated joint engagement would be especially problematic for children with autism and that supported joint engagement would be relatively spared. Compared to DS and TD children, children with autism were less likely to be observed in coordinated joint engagement but did not differ in supported joint engagement. Moreover, DS children did not differ from TD children in supported or coordinated joint engagement, although they were less likely to be engaged in solitary object play.

Our expectations about symbol-infused joint engagement were generally supported. However, counter to our expectation that symbol-infusion might prove especially difficult for the children with Down syndrome, diagnosis per se did not affect the amount of symbol-infusion when 30-month-old children with autism or with Down syndrome were compared with 18-month-old TD children.

Effect of Communicative Context on Joint Engagement

In addition to our primary hypotheses about overall patterns of joint engagement, we also wanted to assess whether children with autism would have particular difficulty during contexts that afforded commenting and children with Down syndrome would find requesting contexts especially challenging. Our hypotheses were partially supported. We conducted analyses of variance, separately for the three diagnostic groups and for time spent unengaged and for each of the three joint engagement variables, with context as the repeated measure. Context strongly affected time unengaged for the AU and DS samples (see Table 3). Children with autism were more often unengaged during commenting than either interacting or requesting, whereas children with Down syndrome were more often unengaged during both requesting and commenting than interacting. However, for all three samples, more time was spent in symbol-infused joint engagement during the commenting than the interacting context, although for the AU sample this difference (4.7) did not reach the Tukey required 4.9. Context did not affect time spent in supported or coordinated joint engagement ($\eta^2 < .09, p > .10$ for all).

Joint Engagement Trajectories

Our second aim was to discern how joint engagement developed during the year we observed children and whether developmental trajectories were different for children with autism and Down syndrome compared to TD children. For the trajectory analyses we computed the slope (unstandardized regression coefficient) for the symbol-infused, supported, and coordinated joint engagement variables for each child over their 3–5 visits (excluding 3 AU children and 1 DS child with only two visits and 2 TD children with only one CDI score). These slopes were then analyzed with diagnostic by language trajectory group analyses of variance (language group was defined by the CDI trajectory; see Method section).

The analyses of variance revealed that for all three variables (symbol-infused, supported, and coordinated joint engagement) mean slopes differed from zero ($\eta^2 = .81, .10, and .20, p < .001, = .002, <.001$), but with some variation. For symbol-infused joint engagement slopes were positive (indicating an increase with age) for 95, 93, and 100% of the AU, DS, and TD children, respectively; these percents did not differ significantly ($\chi^2[2] = 3.67, p = .16$) but all significantly exceeded the 50% expected by chance ($p < .001$) per two-tailed sign test. Supported joint engagement slopes were positive for 75, 43, and 70% of the AU, DS, and TD children; these differed among themselves ($\chi^2[2] = 7.41, p = .025$) and only the AU and TD percents differed from chance, $p < .05$ and $.001$, respectively. Finally, coordinated joint engagement slopes were positive for 60, 64, and 74% of the AU, DS, and TD children; these did not differ significantly ($\chi^2[2] = 1.68, p = .43$) and only the TD percent differed from chance, $p < .001$.

The analyses of variance also revealed that for all three variables mean slopes were affected by language trajectory group (albeit marginal for supported) but no interaction with diagnostic group was significant ($\eta^2 < .09, p > .10$ for
For symbol-infused and supported but not coordinated joint engagement, mean slopes varied by diagnostic group as well. Mean slopes for diagnostic and for language trajectory groups, along with the strength and significance of the main effects, are given in Table 4; Fig. 2 graphs the engagement state variables separately by diagnostic group. Symbol-infused joint engagement showed the greatest development over the year (i.e., had the steepest slopes); its increase was greatest for TD, and for emerging and already verbal, children. Supported joint engagement showed much weaker development: It was likewise highest for TD, and already verbal, children but actually declined a bit for DS, and minimally verbal, children; for all groups, it averaged around 50% (see Fig. 2). Coordinated joint engagement likewise showed relatively weak development over the year. Its slight increase was similar for all three diagnostic groups but at different levels (but a slight decrease for already verbal children); lowest was the AU group, followed by the DS and TD groups, respectively (again, see Fig. 2).

Table 3 Effect of context on symbol-infused joint engagement and unengaged by diagnostic group

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diagnostic group</th>
<th>Context means</th>
<th>Context effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Interacting</td>
<td>Requesting</td>
</tr>
<tr>
<td>Symbol-infused</td>
<td>Autism$_{30}$</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Down$_{30}$</td>
<td>4$_a$</td>
<td>6$_{ab}$</td>
</tr>
<tr>
<td></td>
<td>Typical$_{18}$</td>
<td>4$_a$</td>
<td>9$_b$</td>
</tr>
<tr>
<td>Unengaged</td>
<td>Autism$_{30}$</td>
<td>11$_a$</td>
<td>14$_a$</td>
</tr>
<tr>
<td></td>
<td>Down$_{30}$</td>
<td>4$_a$</td>
<td>8$_b$</td>
</tr>
<tr>
<td></td>
<td>Typical$_{18}$</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

Note: Means are percents. Effect sizes are from repeated measures analyses of variance; $n =$ 23, 29, and 56 for the AU$_{30}$, DS$_{30}$, and TD$_{18}$ samples, respectively. For each diagnostic group, means that did not differ significantly per a Tukey post hoc test, $p < .05$, share a common subscript.

Table 4 Trajectories for joint engagement state variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diagnostic group means</th>
<th>Diagnostic effect</th>
<th>Language group means</th>
<th>Language effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Autism$_{30}$</td>
<td>Down$_{30}$</td>
<td>Typical$_{18}$</td>
<td>$\eta^2$</td>
</tr>
<tr>
<td>Symbol-infused</td>
<td>1.30$_a$</td>
<td>0.95$_a$</td>
<td>3.65$_b$</td>
<td>.54</td>
</tr>
<tr>
<td>Supported</td>
<td>0.27$_{ab}$</td>
<td>-0.08$_a$</td>
<td>0.48$_b$</td>
<td>.07</td>
</tr>
<tr>
<td>Coordinated</td>
<td>0.31$_{b}$</td>
<td>0.38$_{b}$</td>
<td>0.50$_{b}$</td>
<td>.01</td>
</tr>
<tr>
<td>Symbol-infused</td>
<td>0.24$_a$</td>
<td>2.87$_b$</td>
<td>2.04$_a$</td>
<td>.10</td>
</tr>
<tr>
<td>Supported</td>
<td>-0.29$_a$</td>
<td>0.27$_{ab}$</td>
<td>0.73$_{b}$</td>
<td>.06</td>
</tr>
<tr>
<td>Coordinated</td>
<td>0.63$_b$</td>
<td>0.56$_{ab}$</td>
<td>-0.26$_a$</td>
<td>.12</td>
</tr>
</tbody>
</table>

Note: Means are linear slopes (unstandardized regression coefficients) and represent the increase in the joint engagement state variable in percents for each additional month of age. Effect sizes are from diagnostic group by language trajectory group analyses of variance, $N =$ 102; interaction effects ($\eta^2 < .09$, $p > .10$, for all) are not shown. For each variable within diagnostic or language trajectory group, means that share a common subscript did not differ significantly, $p < .05$, per a Tukey post hoc test.

Prediction of Later Language

Our last set of analyses evaluated the association between symbol-infused joint engagement and children’s subsequent receptive and expressive vocabulary. PPVT-III scores from the children’s fifth visit were available for 18 AU, 22 DS, and 53 TD children ($n = 93$); corresponding numbers for EVT scores were 18, 22, and 54 ($n = 94$). For these analyses, we evaluated symbol-infused supported and symbol-infused coordinated joint engagement separately; scores represented cumulative experience, averaged over

---

All) For symbol-infused and supported but not coordinated joint engagement, mean slopes varied by diagnostic group as well. Mean slopes for diagnostic and for language trajectory groups, along with the strength and significance of the main effects, are given in Table 4; Fig. 2 graphs the engagement state variables separately by diagnostic group. Symbol-infused joint engagement showed the greatest development over the year (i.e., had the steepest slopes); its increase was greatest for TD, and for emerging and already verbal, children. Supported joint engagement showed much weaker development: It was likewise highest for TD, and already verbal, children but actually declined a bit for DS, and minimally verbal, children; for all groups, it averaged around 50% (see Fig. 2). Coordinated joint engagement likewise showed relatively weak development over the year. Its slight increase was similar for all three diagnostic groups but at different levels (but a slight decrease for already verbal children); lowest was the AU group, followed by the DS and TD groups, respectively (again, see Fig. 2).
the year’s observations. For each diagnostic group, we regressed PPVT and EVT scores, first on the CDI score from the AU and DS children’s 30-month visit and the TD children’s 18-month visit (to control for initial differences in language ability); second on symbol-infused supported joint engagement (to determine its influence above and beyond initial language), and finally on symbol-infused coordinated joint engagement (to determine whether it made any additional contribution).

Descriptive statistics for the predictor and outcome variables are given in Table 5 and reflect expected patterns. Initial CDI scores were higher for the AU than the DS sample, but neither was significantly different from the TD sample. Symbol-infused supported experience was comparable for the AU and TD samples; the mean for the DS sample was significantly lower. In contrast, symbol-infused coordinated experience was comparable for AU and DS samples; the mean for the TD sample was significantly higher. Correlations between symbol-infused supported and coordinated joint engagement were .74, .33, and .36, \( p < .001, = .13, \text{ and } = .008 \) for AU, DS, and TD samples, respectively. PPVT and EVT scores were lowest for the DS, higher for the AU, and highest for the TD sample. Correlations between PPVT-III and EVT scores were .90, .60, and .75, \( p < .001, = .003, \text{ and } < .001 \) for AU, DS, and TD samples, respectively.

Multiple regression results are given in Table 6. After controlling for initial CDI scores, the amount of supported symbol-infused experience accounted strongly for variability in both PPVT-III and EVT scores in all three diagnostic groups (changes in \( R^2 \) were .23 for PPVT-III AU scores and exceeded .25 for others). Coordinated symbol-infused experience accounted for statistically significant, moderate additional variance only for DS EVT scores; otherwise the unique contribution of coordinated symbol-infused experience was weak and not statistically significant. As an additional check (as in Adamson et al. 2004), we reversed the order, entering supported symbol-infused scores on the last step, which tests the unique additional variance accounted for by supported symbol-infused experience after taking coordinated symbol-infused experience into account. In all cases, even though coordinated symbol-infused effects were weak to moderate, supported symbol-infused experience accounted uniquely for moderate to strong additional variance.

Discussion

The results of this study indicate that autism and Down syndrome often affect a young child’s joint engagement experiences during social interactions with a caregiver. Compared to typically developing peers, children with autism rarely coordinated attention to a shared object and the partner, a deficit that was no less marked in children who had acquired relatively large vocabularies. In contrast, children in the DS group readily shared events with their partners but they were less likely to attend to symbols during these

Table 5 Descriptive statistics for language outcome and predictor variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diagnostic group means</th>
<th>Diagnostic effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Autism</td>
<td>Down</td>
</tr>
<tr>
<td>Initial CDI</td>
<td>113</td>
<td>40</td>
</tr>
<tr>
<td>Supported SI</td>
<td>15</td>
<td>7.2</td>
</tr>
<tr>
<td>Coordinated SI</td>
<td>5.3</td>
<td>5.6</td>
</tr>
<tr>
<td>PPVT-III</td>
<td>73</td>
<td>56</td>
</tr>
<tr>
<td>EVT</td>
<td>78</td>
<td>58</td>
</tr>
</tbody>
</table>

Note: Effect sizes are from diagnostic group analyses of variance; \( n = 18, 22, \text{ and } 55 \) for all AU, DS, and TD variables, respectively, except TD PPVT-III (\( n = 53 \)) and TD EVT (\( n = 54 \)). For each variable, means that share a common subscript did not differ significantly, \( p < .05 \), per a Tukey post hoc test.

Table 6 Predicting language outcome from symbol-infused supported and coordinated joint engagement experience

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Step</th>
<th>Variable added</th>
<th>Autism</th>
<th>Down</th>
<th>Typical</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT-III</td>
<td>1</td>
<td>Initial CDI</td>
<td>.47</td>
<td>.06</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Supported SI</td>
<td>.23</td>
<td>.31</td>
<td>.30</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Coordinated SI</td>
<td>.06</td>
<td>.07</td>
<td>.00</td>
</tr>
<tr>
<td>EVT</td>
<td>1</td>
<td>Initial CDI</td>
<td>.39</td>
<td>.13</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Supported SI</td>
<td>.41</td>
<td>.31</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Coordinated SI</td>
<td>.00</td>
<td>.11</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Coordinated SI</td>
<td>.14</td>
<td>.22</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Supported SI</td>
<td>.28</td>
<td>.20</td>
<td>.31</td>
</tr>
</tbody>
</table>

Note: SI = symbol-infused. Changes in \( R^2 \) show two hierarchic regression orders, first for supported before coordinated after initial CDI, and then the reverse. \( n = 18, 22, \text{ and } 55 \) for all AU, DS, and TD analyses, respectively, except TD PPVT-III (\( n = 53 \)) and TD EVT (\( n = 54 \)).
periods. It is noteworthy that the influence of diagnosis on the form of joint engagement was remarkably stable over time. The low amount of coordinated joint engagement in children with autism and of symbol-infusion in children with Down syndrome was evident both when comparisons were made at 30 months, an age when both coordinating attention and symbol use are typically consolidated, and in comparisons when children had comparable expressive vocabularies. Furthermore, when the developmental course was plotted for a year (beginning when children with autism and with Down syndrome averaged 31 and 30 months of age, respectively, and TD children were 18 months of age, so that language abilities were comparable), coordinated joint engagement did not emerge in autism and symbol-infused joint engagement developed much less strongly for the children with Down syndrome.

These findings elaborate prior reports of early joint attention skill deficits that characterize autism and of the expressive language problems that challenge children with Down syndrome by providing a rare view of the development of joint engagement and symbol use during the optimizing setting of uninterrupted play with a caregiver. The Communication Play Protocol provided a standard premise that allowed us to observe the child in different communicative contexts and over time. Children were almost always engaged with objects and/or people during the Play; even children with autism who were more likely unengaged than typically developing children or children with Down syndrome, were on average engaged 85% of the 30 min long observation period. Nevertheless, some communicative contexts were more difficult than others for toddlers with autism, who were most often unengaged during commenting contexts, and for toddlers with Down syndrome, who were most often engaged during interacting contexts. These context differences, which were not apparent in the typically developing sample, suggest that it is important to consider not only how a disorder may impact forms of joint engagement but also how it may impact a child’s willingness to become engaged in interactions that focus on specific functions such as commenting and requesting.

This view supports three related conclusions. First, there are striking variations in the developmental path of symbol-infused joint engagement. Second, along some of these paths, coordinated joint engagement may be neither a necessary nor a sufficient step towards the introduction of symbols into parent–child interactions. Finally, and most intriguingly, across all of the groups in this study, periods of symbol-infused supported joint engagement during caregiver-child interactions may provide an especially facilitative context for early language learning.

With regards to developmental paths, we found two variations from the path taken by typically developing toddlers where the consolidation of a triadic arrangement of attention between self, partner, and shared objects occurred before symbols infuse social interactions (Adamson et al. 2004). The first variation was marked by the achievement of a triadic nonverbal communication structure that was not followed rapidly by the emergence of symbol-infused joint engagement. A substantial portion of the group of young children with Down syndrome we observed had not started to infuse symbols regularly into joint engagement by age 2½, although all did do so at least minimally at some point during our year-long observation period. A similar pattern of coordinated joint engagement without symbol-infusion has been observed in a study of deaf 22 month-olds interacting with their hearing mothers (Prezbindowski et al. 1998) and in an on-going study of young children with severe speech and developmental delays prior to effective parent-implemented language intervention (Romski et al. 2008). Taken together, these findings suggest that a certain level of vocabulary, and perhaps more generally “symbol-mindedness” (DeLoache 2002), are needed for a child to actively infuse symbols into joint engagement. Not surprisingly, infusing symbols into joint engagement depended on current verbal skill. Minimal verbal children, be they typically developing but late talking 18-month olds (Adamson et al. 2004) or 30 month old children with autism or with Down syndrome, were rarely observed in symbol-infused engagement states. Further, children who became verbal during our year-long study also became increasingly able to infuse symbols into joint engagement. But the abstraction of this pattern of delay also underscores the need for research that more fully explores the specific barriers to symbolization faced by young children with Down syndrome (Yoder and Warren 2004) and other children who have mastered the rudiments of coordinated joint engagement.

The second variation in the typical developmental path involves a movement towards increased symbol-infusion without the consolidation of coordinated joint engagement either before or during the emergence of symbol-infused joint engagement. This pattern was more likely to characterize children with autism (for whom symbol-infused joint engagement occurred more often than coordinated joint engagement in 52% of the participants) than in either the typically developing or Down syndrome samples (for whom symbol-infused joint engagement occurred more often than coordinated joint engagement in only 17 and 10% of the participants, respectively). There are a myriad of reasons why young children with autism might have difficulty sustaining periods of coordinated joint engagement even when interacting with a caregiver who was trying to facilitate communication. The child may appear uninterested in the partner (Osterling and Dawson 1994), or fail to orient to her (Dawson et al. 1998; Leekam and Moore 2001) and be affectively unresponsive (Joseph and Tager-Flusberg 1997; Hobson 1993). Highly restricted or
idiosyncratic object interests may also make it difficult to locate a topic for sustained shared attention (Williams et al. 1999) and conversation (Watson 1998). Moreover, the child might have difficulty integrating elements during the rapid flow of a social interaction, and thereby fail to alternate gaze between a partner’s face and an object (e.g., Charman 2004), coordinate an emotional response to an object and gaze to a partner, or produce a well-timed declarative point (Travis and Sigman 2001).

Alongside these findings of difficulty with coordinated and with symbol-infused joint engagement, it is particularly noteworthy that young children with autism and with Down syndrome were often able to sustain periods of supported joint engagement at a rate comparable to that observed in typically developing 18- and 30-month old children. This rate was usually substantial; on average, children spent approximately half of each 30 min communication play in supported joint engagement. Moreover, in all of the three groups, supported joint engagement was as likely as coordinated joint engagement to be infused with symbols. Thus even children with autism who found coordinated joint engagement problematic could enter a sphere of shared focus in which partners’ actions blend with their own. Furthermore, even children with Down syndrome, who were less likely than others to engage in solitary object play, often shared objects without simultaneously coordinating attention to their partner.

Certainly much of the credit for negotiating supported engagement states accrues to the caregivers, especially when their partner is a child with autism who rarely acknowledges bids for joint attention and who might even actively try to thwart them (Adamson et al. 2001). Although there are yet few studies that focus on the caregiver’s contribution to children’s engagement states, there is evidence that parents interacting with children with autism or with other developmental disorders may synchronize their actions with the child’s actions at the same rate as parents interacting with a typically developing child (Siller and Sigman 2002), and that they may modify their actions in ways that increase the salience of objects by, for example, supplementing words and pointing gestures with more literal acts such as banging and waving (Baranek 1999; McArthur and Adamson 1996). Moreover, evidence is mounting that parents may implement interventions that promote joint attention in toddlers with autism (e.g., Schertz and Odom 2007).

Nevertheless, it is also important to appreciate the child’s contribution. For a period to be coded supported joint engagement, a child, as well as his or her caregiver, had to be actively involved in object or event sharing. Given our coding scheme’s criteria, this meant that the child was not merely watching the caregiver (which would have been coded as on-looking) and that he or she sustained attention to a shared topic for at least 3 s (fleeting interest would have been coded unengaged). Thus, supported joint engagement indicates that the child had at least a nascent capacity to move beyond a singular focus on objects to share objects in a way that incorporates a partner’s contribution, however implicit and unappreciated and however dependent on extraordinary scaffolding.

Most intriguingly, in all three groups, variations in how often children were observed in symbol-infused supported joint engagement predicted the growth of their receptive and expressive vocabularies, after controlling for initial language level. These findings about the significance of periods of symbol-infused supported joint engagement converge well with mounting evidence that periods of time when a language learning child focuses primarily on an event and its symbolic representation, rather than on the full triad of partner, object, and symbol, may be especially conducive to new word learning. For example, Bloom (1993; Bloom and Tinker 2001) presents compelling evidence that arrangements that lessen the cognitive and affective demands of interpersonal communication and heighten the relation between symbol and referent may help a child focus on the difficult task of acquiring language. Early language acquisition seems to be facilitated when mothers follow their child’s lead rather than direct their child’s attention (Tomasello and Farrar 1986). Moreover, children’s ability to respond to joint attention bids, rather than their ability to initiate them, has often been found to predict variations in early language learning in young children, including within samples of children with autism (Bono et al. 2004; Sigman and Ruskin 1999; Siller and Sigman 2002; Sullivan et al. 2007; cf. McDuffie et al. 2005).

By the end of the study, all of the typically developing children, 78% of the children with autism, and 79% of the children with Down syndrome had acquired at least a 50 word expressive vocabulary. These rates, especially for children with autism, are relatively high compared to older estimates (Tager-Flusberg 1994; cf. Lord et al. 2004). Since the children with autism participated in early intervention programs, their success may underscore the promise of early detection and intervention (Landa 2007; Wetherby and Woods 2006), although it may also reflect changes in diagnostic criteria over time. We hasten to add that our findings do not indicate that periods of symbol-infused coordinated joint engagement are unimportant to language acquisition. Focusing only on the symbol-referent relationship during supported joint engagement may not provide a young child with the information about meaning that is gained when toddlers monitor their partner’s attention and intentions (Baldwin 1995). Periods of coordinated joint engagement may provide a particularly rich context for toddlers to learn about theory of mind.
specifically (Nelson et al. in press) and about how to participate in the decontextualized, connected conversations that typically emerge during the preschool years (Adamson and Bakeman 2006; Nelson 1996). Thus even if children with autism acquire language, they may continue to experience problems using it appropriately in conversations (Tager-Flusberg 1994), and they may display a peculiar fascination with symbol systems such as letters, numbers, and names outside the sphere of joint engagement (Frith 1989).

These findings highlight the importance of looking beyond periods of coordinated joint engagement to study the way additional arrangements of attention might allow children to engage in language-facilitating interactions, including children with impaired joint attention skills and those who find the step into the symbolic sphere especially challenging. This contention affords with the growing appreciation (articulated well in Akhtar 2005; Hoff 2005) that although all typically developing children develop joint attention skills and language, not all cultures embed early object exploration or first words in a child-centered social context. Moreover, these findings encourage us to continue to study the transaction between children and their partners at various points along the path of communication development in hopes of understanding more fully how others may provide scaffolds that facilitate a child’s symbol formation and use.

Acknowledgments This research was supported by grants from the National Institute of Child Health and Human Development (R01HD35612) and from the Research Enhancement Program of Georgia State University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Child Health and Human Development or the National Institutes of Health. Portions of the study were presented at the Society for Research in Child Development, Atlanta, 2005, and at the International Conference on Infant Studies, Kyoto, 2006. The authors thank Barbara Dunbar, Pamela K. Rutherford, Janis Sayre, and Kimberly M. McMillan for their many contributions to this project and Diana Robins for her helpful comments. In addition we gratefully acknowledge Alicia Brady, M. Janelle Cambron, Jesse Centrella, Susan Kupferberg, Ginair McKerrow, P. Brooke Nelson, Lauren Pierre, Jana Pruett, Rodney Teague, Melissa Traver, Anjali Vasudeva, and Leslie Wade who coded the corpus.

References


Springer


