Many researchers look at language as a purely formal object. In this approach, the syntax of language is innate, and even the acquisition of individual words is underlain by innate word-learning constraints. Language is thought to be a cognitive module distinct from other cognitive and social capacities (Chomsky, 1986; Pinker, 1995). In contrast, the social-pragmatic approach to language and its acquisition views language very differently (Bruner, 1983; Nelson, 1985; Tomasello, 1992). In this view, each of the world’s natural languages has its own set of communicative conventions, in the form of linguistic symbols created over thousands of years, by means of which its speakers attempt to influence the interest, attention, knowledge, and behavior of other members of their speech communities. This is not to deny that there are universals in the way symbols are created, learned, and used across languages. Instead of being innate and language-specific, however, these universals are universals both in the way human beings experience the world and in the ways they interact and communicate with one another socially.

In the social-pragmatic approach to language acquisition, the focus is on both the structured social world into which the child is born and the child’s capacities for tuning into and participating in that structured social world (Tomasello, 1992). In this view, young children are not engaged in a reflective cognitive task in which they are attempting to make correct “mappings” of word to word based on adult “input”; rather, they are engaged in social interactions in which they are attempting to understand and interpret adult communicative intentions so as to make sense of the current situation (Nelson, 1985). When attempting to comprehend adult use of novel linguistic symbols, children use all kinds of interpretive strategies based on the pragmatic assumption that adult linguistic symbols are somehow relevant to the ongoing social inter-
action (Bloom, 1993; Bruner, 1983; Sperber & Wilson, 1986). The child who knows that his mother wishes him to eat his peas (because she is holding them up to his mouth and gesturing) assumes that her utterance is relevant to that intention, and this understanding of communicative intentions is what guides that child's interpretations of any novel language in the situation. The child may then learn to produce the same symbols when he wishes for others to experience a situation in the same way, thus entering into the world of bidirectionally (intersubjectively) understood linguistic symbols (Tomasello, 1995).

In the social-pragmatic view, then, children acquire linguistic symbols as an integral part of their social interactions with adults in much the same way that they learn many other cultural conventions. Tomasello, Kruger, and Ratner (1993) have called this process cultural learning. In this chapter we spell out this view of language acquisition in more detail, focusing first on how children "get off the ground" in language acquisition via processes of joint attentional interaction, then on how they progress in word learning in the second year of life, and then on the social-cognitive skills on which language acquisition depends. Finally, we apply this general theoretical approach to the acquisition of language by children with autism. The language difficulties of these children are well known, but we believe they may be better understood when they are seen in the light of other difficulties these children have with social and communicative activities in general.

**JOINT ATTENTION AND EARLY LANGUAGE**

The problem of how children learn linguistic symbols was first clearly articulated by Wittgenstein (1953): How can a child learn a word when nonlinguistic procedures can unambiguously illustrate its reference? Wittgenstein noted that even an ostensive definition—the seemingly simplest case of language acquisition in which one person "shows" another what a word means—is problematic because it assumes that both teacher and learner know what "showing" is and precisely how it serves to pick out individual referents in some language-independent way. The point was crystallized by Quine (1960) in his parable of a native who utters the expression "Gavagai" and "shows" a foreigner the intended referent by pointing out a salient event as it unfolds. Given the stipulation that native and foreigner have no way to establish a common view of the event nonlinguistically, however, there is basically no way that the foreigner can know whether the native's novel expression is being used to refer to the event, to some participant in the event, to some part of the participant's body, to the color of the participant's hair, or to any of an infinite number of aspects of the situation. This is the basic problem of referential indeterminacy.

Bruner (1975, 1983) addressed this problem and gave a basically Wittgensteinian answer: The child acquires the conventional use of a word by learning to participate in a form of life that he or she understands first non-linguistically, so that adults' language can be grounded in shared experiences whose social significance he or she already appreciates. One key component of this process is a child who can understand adults as intentional beings well enough to share attention with them in some shared activity (we discuss this later). There is another component, however, when what is at issue are things cultural and conventional. To learn a conventional symbol the child must live in a world that has structured social activities in which he or she can participate and whose structure he or she understands. These structured social activities often involve the recurrence of the same general activity on a regular basis so that the child can come to comprehend how the activity works, how the various social roles in it function, and so forth. And, of course, it must be the case that the adult uses a novel linguistic symbol in a way that the child can comprehend within that shared activity.

The basic idea may be illustrated by imagining an adult attempting to learn a foreign language. In one scenario, one can imagine the adult trying to learn from interactions in which a native speaker simply starts and stops talking in the foreign tongue—out of nowhere, so to speak. It is very unlikely that in such situations the learner would acquire the conventional use of any forms from the foreign language at all. In another scenario, we may imagine that the learner enters a store or a train station and begins interacting with the storekeeper or ticket-seller in some meaningful ways. In these situations it is highly likely that some new language will be learned because both interacants share an understanding of one another's interactive goals in terms of purchasing and selling items, and so it is very likely that in many cases the learner can infer the native's communicative intentions independently of language. As the two interact in searching for objects, exchanging objects, exchanging money, and so forth, the learner may actually acquire some new language if the native speaker uses that new language in some way that suggests to the learner some reason for making that utterance at that time. This may happen, for instance, when the native speaker utters a novel word while holding out his or her hand from behind the cash register after the learner has acquired an object from the shelf. In such cases the learner makes an abductive inference of the following type: If that novel word meant X, then it would be relevant to our ongoing intentional interactions (Nelson, 1996; Sperber & Wilson, 1986).

A variety of studies have shown that after children have begun progressing in the process of language acquisition, they learn new words best in joint attentional interactions, in which children and adults coordinate their attention to each other and an object of mutual interest. This coordination is evident when children engage in referential looking, or gaze alternation between the object and the adult's face (e.g., Bakeman & Adamson, 1984; Trevarthen & Hubley, 1978). These joint attentional interactions often take place during routine situations, such as bathing, feeding, diaper changing, book reading, and traveling in the car, that are recurrent activities in the child's daily experience.
These activities are in many ways analogous to the store scenario in that children are likely to understand their own and the adult's goals in the situation, and this understanding may enable children to infer the relevance of the adult's language to those goals. As we discuss later, children with autism, in contrast, have difficulty engaging in joint attentional interactions and thus may experience language more often than not in situations analogous to the first scenario that we mentioned, in which the language comes from "out of nowhere."

One such study was conducted by Tomasello and Todd (1983), who documented that mother-child dyads who spent more time in joint attentional activities during the 12- to 18-month period had children with larger vocabularies at 18 months of age (see also Smith, Adamson, & Bakeman, 1988, and Tomasello, Manne, & Kruger, 1986). The timing of adults' language in these activities is important as well. In two correlational and experimental studies, Tomasello and Farrar (1986) found that mothers who used language to follow into their child's attention (i.e., to talk about an object that was already the focus of the child's interest and attention) had children with larger vocabularies than mothers who used their language in an attempt to direct the child's attention to something new (see also Akhtar, Dunham, & Dunham, 1991; Dunham, Dunham, & Curwin, 1993).

These studies demonstrate the importance of joint attentional activities and adult language that is sensitive to the child's current focus of attention in the language acquisition of children 18 months of age or older—an age by which many children have already become quite adept at word learning. A remaining question thus is whether these relations hold at even younger ages, when language skills are beginning to emerge just after the child's first birthday. This is more than just a question of whether relations can be demonstrated earlier; it also involves the issue of whether children's earliest language skills emerge out of the prelinguistic joint attentional activities just described or whether they arise from some other source.

The study of most direct relevance to this question is that of Carpenter, Nagell, and Tomasello (1998), who followed 24 infants longitudinally, at monthly intervals, from 9 to 15 months of age. At each monthly session infants interacted naturally with their mothers for 10 minutes, and from these interactions the following measures were computed: 1) the amount of time each mother-infant dyad spent in joint attentional engagement and 2) the percentage of maternal utterances that followed into the infants' focus of attention. At these same monthly sessions the mothers also reported, via a structured checklist, each of the words that their infants had mastered in comprehension and production (they also reported the different gestures their infants produced).

Two main findings emerged. First, mother-infant dyads who spent more time in joint attentional engagement at 12 months of age had infants who used more gestures and comprehended more language in the months immediately following. A similar relation emerged a few months later between joint attentional engagement and language production. Second, mothers who followed into their infants' attentional focus with referential words at 12 months of age had infants with larger comprehension vocabularies in the months immediately following (with a similar relation to language production again showing up a bit later). When both of these variables—time in joint attentional engagement and mother's follow-in language—were used together in multiple regression equations, more than half of the variance in infants' language comprehension and production was predicted at several points during the period from 12 to 15 months of age, with each variable accounting for significant amounts of unique variance.

Several other findings of this study are important. First, a number of other social-cognitive measures (e.g., gaze following, imitative learning, gestural communication) of these infants were taken as well. All 24 of the infants had engaged meaningfully in these other joint attentional activities before they actually began producing conventional language. This fact, coupled with the fact that the emergence of linguistic skills was correlated with mother-infant joint attentional interactions, suggests that language skills indeed may be seen as emerging out of nonlinguistic joint attentional activities. In the most radical view of this process, children's nonlinguistic joint attentional activities may be said to come to be supplemented by linguistically mediated joint attentional activities. Second, two measures of infants' non-social-cognitive development (involving their knowledge of objects and space) were also taken. These mostly emerged in an uncorrelated fashion with language and the other joint attentional activities. This finding provides evidence that the relation between the emergence of joint attentional engagement and language is not just the result of some generalized developmental advance (e.g., in attention span). The relations among these social behaviors are much more intimately linked than that.

Finally, it is also noteworthy that the relation between maternal follow-in language and infants' language became weaker as infants grew older. This is an intriguing finding because it suggests the possibility—explored and supported later in this chapter—that mothers' use of language that follows into the infants' attentional focus is a kind of scaffolding for early language, which helps infants who are just getting started to discern the mother's communicative intentions and so to enter into a state of joint attentional focus. This kind of scaffolding, however, may not be necessary as the child grows older and becomes more skillful in determining communicative intentions in less child-friendly linguistic interactions. Something similar, but different, may occur with joint attentional activities as older children learn to establish and maintain joint attentional interactions on the basis of language itself, for example, using language to jointly attend with an adult to some past event that is no longer perceptually present (Tomasello, 1988).
The clear finding of the Carpenter, Nagell, and Tomasello (1998) study—which confirms the findings of similar studies of slightly older children—is that children’s emerging ability to engage in nonlinguistically mediated joint attentional activities with adults at approximately 1 year of age is integrally related to their emerging linguistic skills. This finding is important because it demonstrates that the well-known age correspondence between joint attentional skills and language (which emerge, respectively, in the months before and after the child’s first birthday) is not a coincidence. This finding presents an immediate and serious problem for theories of early language acquisition that do not focus on the social dimension of the process. For theories that focus primarily on the cognitive dimensions of word learning (e.g., Markman, 1989), or on the learning processes involved (Smith, Jones, & Landau, 1996), the question is why does language acquisition begin when it does? Why does it begin directly on the heels of the emergence of joint attentional skills? Any answer that invokes non-social-cognitive or learning processes—for example, that children at this age for the first time become able to conceptualize or to learn new sorts of things in general—must then answer the question of why early language emerges in a correlated fashion with these nonlinguistic cognitive and social-interactive skills. To our knowledge, none of the existing theories of early language acquisition—other than the social-pragmatic theory as espoused by Bruner (1983), Nelson (1985), Tomasello (1992, in press), and others—can account for these findings.

LEARNING LANGUAGE IN THE FLOW OF SOCIAL INTERACTION

A number of studies have thus established that very young language learners benefit from adult language models that follow into their already-established focus of attention on an object. Children, however, soon come to be able to determine adults’ communicative intentions in a much wider variety of communicative contexts in which they have to do much more social-cognitive work. It does happen with some frequency that in Western middle-class culture, an adult holds up or points to an object while telling the child its name. The social dimensions of this process are manifest: The child must somehow determine the focus of the adult’s attention. In this case, though, the social-cognitive task would seem to be at least relatively straightforward because such things as following gaze direction are so basic for infants.

It turns out, however, that in many cultures of the world adults do not engage in this kind of naming game with young children (Brown, in press). Moreover, even in Western middle-class culture, adults do not frequently use this naming game with words other than object labels; for example, they use verbs most often to regulate or anticipate children’s behavior, not to name actions for them (Tomasello & Kruger, 1992). It would seem bizarre indeed if an adult were to exclaim to a child: “Look, this is an instance of putting. Look at me putting this here.” Instead children typically hear common verbs as an adult directs their behavior in such utterances as “Put your toys away” while pointing to the toy box. It is clear that in such cases, the social-pragmatic cues that might indicate the adult’s intended referent for the child are much more subtle, complex, and variated than in the ostensive context. Indeed, they even change in fundamental ways from situation to situation: The adult requests that the child eat peas by directing the spoon at the child’s face, requests that the child give something by holding out a hand, and requests the putting away of toys by pointing to the destination desired. Thus, there is no standardized “original naming game” for verbs and many other types of early words (e.g., prepositions) as there is for object labels for some children (Tomasello, 1995b).

A number of studies have demonstrated experimentally that children can indeed learn new words in a variety of fairly complex social-interactive situations. Children learn new words not just when adults stop and name objects for them but also in the ongoing flow of social interaction in which both they and the adult are trying to do things. In none of these cases can the child count on the adult following into the child’s already-established focus of attention, but rather the child must adapt to the adult’s focus of attention. For example, Baldwin (1993a, 1993b) taught 19-month-old children new words in two new situations. In one situation the adult followed into the child’s focus of attention, and, as in other studies, children learned the new word quite well. Children, however, also learned new words in a situation in which the adult labeled an object the child was not already looking at, thus requiring the child to look up and then determine the adult’s attentional focus.

In addition, Tomasello and colleagues (see Tomasello, in press, for a review) have conducted a series of studies that demonstrate the same point even more dramatically. The basic idea in all cases is to set up situations in which adults talk to children as they engage in various games, with novel words being introduced as naturally as possible into the ongoing flow of the game. In all cases multiple potential referents are available, that is, there are multiple novel referents for which the child has no existing name. Various social-pragmatic cues to the adult’s intended referent are provided in different studies to see whether children are sensitive to them. The studies are designed so that neither the adult’s gaze direction nor any of the well-known word-learning constraints that various investigators have proposed (e.g., whole object, mutual exclusivity, syntactic bootstrapping; Markman, 1989) will be helpful to the child in distinguishing among possible referents. In all these studies the children ranged from 18 to 24 months of age, and in all cases the majority of children learned the novel words in comprehension, production, or both. To give a feel for the kinds of situations in which children successfully read the adult’s (E’s) communicative intentions and so learn the new word, we summarize some different situations here. In each case, the original study gives all of the details of
control conditions and so forth (there is also an overall summary with many details of these and other similar studies in Tomasello, in press).

- E announced her intention to “find the toma” and then searched in a row of buckets containing novel objects, rejecting some objects by scowling and replacing them until she found the one she wanted (indicated by a smile and the termination of search). Children learned the new word for the object that E’s smile indicated was the one intended, no matter how many rejected objects intervened in the search (Tomasello & Barton, 1994).

- E had the child find four different objects in four different hiding places, one of which was a distinctive toy barn. Once the child had learned which objects went with which places, E announced her intention to “find the gazer.” She then went to the toy barn, but it turned out to be “locked.” She frowned and proceeded to another hiding place to find another toy. Later, when children were asked to pick out the gazer, they did so even though they had never seen the object after they had heard the new word (Study 1 of Akhtar & Tomasello, 1996; see also Study 2 of Akhtar & Tomasello, 1996, in which a similar procedure was used for verb learning).

- E announced her intention to “dax Mickey Mouse” and then proceeded to perform one action accidentally and another intentionally (or sometimes in the reverse order). Children learned the word for the intentional, not the accidental, action regardless of which came first in the sequence (Tomasello & Barton, 1994).

- A child, her mother, and E played with three novel objects; then the mother left the room, and the child and E played with a fourth object. When the mother returned, she looked at the four objects together and exclaimed “Oh look! A modi! A modi!” Understanding that the mother was likely to be excited about the object she was seeing for the first time, children learned the new word for that object (Akhtar, Carpenter, & Tomasello, 1996).

The point of these studies was to investigate something of the range of situations in which children might be able to discern the adult’s communicative intentions and to learn the new word. Although any one of these studies might be explained in other ways (e.g., Samuelson & Smith, 1998), in our view when they are considered as a group, the most plausible explanation is that by the time children are 18–24 months of age, they have a deep and flexible understanding of other individuals as intentional beings. Thus, they are quite skillful at determining the adult’s communicative intentions in a wide variety of novel communicative situations. Children’s assumption that the adult’s language is relevant to their ongoing social and instrumental activities is simply the natural expression of this intentional understanding.

In all cases, no matter what the situation, children must be able to do two things to acquire a communicative symbol. First, they must be able to deter-

mine adults’ communicative intentions. This involves understanding the adult’s intentions toward my attention, that is, understanding what the adult wants me to focus on. This is arguably a more difficult task than determining the adult’s intentions toward, say, a ball, which are much more straightforward. Understanding communicative intentions involves a kind of intentional embedding: understanding his or her intentions toward my attention (Sperber & Wilson, 1986). This understanding allows children to begin to comprehend novel language. Second, to produce conventional linguistic symbols, children must engage in a process of role-reversal imitation in which they use the newly acquired symbol toward the adult in the same way and for the same communicative purpose that the adult used it toward them. Again this process of role-reversal imitation is more complex than imitating, say, the adult’s action on a ball. Imitatively learning a communicative convention means imagining myself in the adult’s role and enacting that action. Thus, learning a new word requires both a special form of the understanding of intentions—understanding communicative intentions—and a special form of cultural learning—role-reversal imitation.

Children get much better at this cultural learning process during the second and third years of life. Initially, at 1 year of age, they are able to learn new language mostly in highly repetitive and predictable social interactions in which the adult follows their attentional focus. As children become more skilled at determining adult communicative intentions in a wider variety of interactive situations, however, highly structured formats with highly sensitive adults become less crucial to the process. The child must establish joint attention in more active ways by determining the adult’s attentional focus in a variety of sociocognitive contexts. Of possible relevance to this account is the finding that some children acquire their native language in cultures in which there is little of the heavy scaffolding and attentional sensitivity that characterize many Western middle-class families. Although quantitative studies have yet to be done, by some accounts these children seldom acquire large numbers of words before their second birthdays (de Leone, personal communication, 1998). If this is true, the implication is that these children acquire the vast majority of their language only after they are able to be more active in determining adult communicative intentions within the flow of ongoing social interaction—as in the verb learning of Western middle-class children and in the experimental studies cited previously.

**SOCIAL-COGNITIVE FOUNDATIONS**

The social-cognitive foundation for language learning is thus the same, with a few twists, as the social-cognitive foundation for children’s other joint attentional activities that first emerge around 1 year of age (e.g., gaze following, social referencing, imitative learning of instrumental actions, perception
of intentional action). All of these skills and activities emerge in rough developmental synchrony because all of them reflect the infant’s emerging ability to understand other individuals as intentional agents whose attention, emotion, and behavior to outside objects may be actively followed into and shared (Carpenter, Nagell, & Tomasello, 1998; Tomasello, 1995a). This social-cognitive revolution at the infant’s first birthday sets the stage for the second year of life in which children culturally learn—imaginatively—the use of all kinds of tools and artifacts, with linguistic symbols being one example. This imaginative learning is not just a mimicking of adult movements and is not just a reproducing of interesting environmental effects by whatever means imaginable; instead it is a reproduction of the adult’s intentional relations to the world. In imaginative learning, the child perceives the adult’s overt actions as composed of both a goal and a means for attaining that goal and then the child actively chooses the adult’s means of goal attainment in contrast to others the child might have chosen. The clearest indications of this type of intentional understanding were shown by Meltzoff (1995), who found that 18-month-old infants imaginatively learned actions that adults intended to perform on objects, even when the adult was unsuccessful in performing them, and by Carpenter, Akhtar, and Tomasello (1998), who found that 14- to 18-month-old infants imitated adults’ intentional actions on objects and ignored their accidental actions.

Our view is thus that the acquisition of linguistic symbols begins during this same developmental period quite simply because comprehending and producing language relies on the same basic understanding of individuals as intentional agents as do all of the other social-cognitive and cultural learning skills that emerge at this same age. Language tends to follow the emergence of the more straightforward expressions of joint attention skills such as gaze following and imitation because it requires a special application of these skills both to understand the special form of intentions known as communicative intentions and to engage in the special form of imaginative learning known as role-reversal imitation. Thus, we believe that disruptions to the foundational joint attentional skills that normally emerge at 1 year of age will have dire consequences for many aspects of children’s language acquisition and use.

CHILDREN WITH AUTISM

Children with autism grow up in the same structured social world as do typically developing children. Despite early proposals to the contrary, there is no evidence that children with autism receive less (or more) social stimulation or worse (or different) parenting than other children (see Cantwell, Baker, & Rutter, 1978b, for a review, and Sigman & Capps, 1997, for discussion). Yet many children with autism never acquire any productive language, and the language of those who do acquire some language is different in many specific ways from that of typical children (this is discussed later in this chapter; see also Frith, 1989, for a review). Consequently, according to the social-pragmatic approach to language acquisition, if these children’s structured social worlds are the same, then there must be something atypical about the ability of children with autism to tune into the social world in which they live (via joint attention), to understand the communicative intentions of those who speak to them, and/or to engage in role-reversal imitation of linguistic symbols. In the remainder of this chapter we review briefly the abilities of children with autism in each of these areas. We then examine studies that have investigated the relations between these abilities and children’s language, and we discuss the implications of deficits in each of these areas for other aspects of language development.

Tuning into the Social World

From very early in life, typical infants are especially attuned to social stimuli (e.g., Morton & Johnson, 1991). Although few studies have been done on very young children with autism because diagnosis is rarely made before age 3 years or so, there is some evidence that these children show differences in orientation to social stimuli compared with both typical children and children with developmental delays as early as 1 year of age. For example, children with autism do not orient to certain speech sounds—their mother’s voice (Klin, 1991) or their own names (Dawson, Meltzoff, & Osterling, 1995; Osterling & Dawson, 1994)—as readily as do other children. (See Chapter 2 for further discussion of early characteristics.) They thus may miss many opportunities to learn new language because they may not attend to much of the speech directed to them.

Similarly, children with autism do not tune into other people in the more active ways that typical children do. Whereas children with autism are social to some degree—for example, they participate in social interactions and games (Dawson, Hill, Spencer, Galpert, & Watson, 1990; Mundy, Sigman, & Kasari, 1990; Sigman, Mund, Sherman, & Ungerer, 1986) and may form secure attachments to caregivers (Capps, Sigman, & Mundy, 1994; Rogers, Ozonoff, & Maslin-Cole, 1991)—these children nevertheless interact with others in fundamentally different ways from other children. Most importantly, children with autism generally do not actively share interest and attention with others, nor do they respond to others’ bids to share interest and attention with them. Many studies (see Charman, in press, for a review) have found deficits in a variety of joint attentional skills: referential looking (i.e., gaze alternation between object and adult; this is the basis of the joint attentional interactions discussed previously; Charman et al., 1997; Lewy & Dawson, 1993; Mundy, Sigman, Ungerer, & Sherman, 1986; Wetherby, Prizant, & Hutchinson, 1998), declarative pointing and showing (Baron-Cohen, 1989; Loveland & Landry, 1986; Sigman et al., 1986; Wetherby et al., 1998; Wetherby & Prutting, 1984), looking
where others look and point (Baron-Cohen, 1989; Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997), and social referencing (Sigman, Kasari, Kwon, & Yimiyi, 1992).

These deficits are not a result of an inability to perform the specific behaviors involved. That is, although children with autism show no differences in amount of overall gaze toward the faces of other people compared with other children (Dawson et al., 1990), they make eye contact in different situations and in different ways. For example, children with autism look at adults' faces as often as other children during face-to-face social games (Sigman et al., 1986). When play involves objects, however, children with autism do not usually alternate gaze between objects and the adult's face (i.e., engage in joint attention). In addition, when they do happen to look at the adult's face, they neither follow the adult's gaze nor use the adult's facial expressions to influence their own behavior to new objects (social referencing); they also do not combine positive affective expressions with their gaze in the way typical children do (Dawson et al., 1990). Similarly, although children with autism often gesture to objects imperatively to request objects or actions from them (often with accompanying gaze alternation between the object they are gesturing about and the other's face), they very rarely gesture declaratively, simply to indicate an object in order to share something about it with the other person (Baron-Cohen, 1989; Curcio, 1978; Wetherby et al., 1998; Wetherby & Prutting, 1984).

This deficit in joint attentional skills is evident very early in development—even before diagnosis is made (Baron-Cohen et al., 1996; Osterling & Dawson, 1994)—and persists throughout childhood (Baron-Cohen, 1989; Landry & Loveland, 1988), although the specific joint attentional deficits present may change with development (Mundy, Sigman, & Kasari, 1994). It reliably differentiates children with autism from children with various other developmental delays (Mundy et al., 1986; Mundy et al., 1990; Wetherby et al., 1998). It is an open question at this point whether lack of joint attention is itself a primary deficit in autism or whether it is instead a consequence of earlier-emerging, more basic deficits such as difficulties with interpersonal relatedness or affective sharing (Hobson, 1993; Mundy & Sigman, 1989) or shifting attention between two stimuli (Courchesne et al., 1994; see also Chapter 8). Either way, the result is that children are prevented from following into and directing others' attention and interest to objects in their shared world. As we discuss later, this difficulty has dire consequences for these children's acquisition of language.

Communicative Intentions and Role-Reversal Imitation

Because joint attention is a precursor of understanding others' communicative intentions and engaging in role-reversal imitation, these latter skills also are likely to be impaired in children with autism. There is very little direct evidence, however, concerning either of these skills in these children. We are aware of no studies of these children's understanding of others' communicative intentions in role-playing or similar situations, similar to Tomasello and colleagues (e.g., Tomasello & Barton, 1994) studies of typically developing children. For example, there is some evidence that suggests that children with autism may have difficulty understanding others' behavioral intentions (Phillips, Baron-Cohen, & Rutter, 1992, 1998); if this is true, then it is likely that this difficulty extends to understanding others' communicative intentions as well.

Similarly, very few studies of children with autism have included tasks designed to measure children's ability to reverse roles when imitating. These children generally show impaired performance on traditional imitation tasks (see Smith & Bryson, 1994, for a review), so the more difficult role-reversal imitation may be even more problematic for them. There is suggestive evidence that this may be the case. Ohta (1987), for example, tested relatively high-functioning children with autism (with a mean chronological age of 10 years) on several gesture imitation tasks. On some of these tests, children with autism sometimes made a distinctive type of error that other children rarely made. For instance, when the modeled action was “waving with the open palm facing the hand,” some children with autism opened their own palm facing themselves (as opposed to palm facing the experimenter); that is, they reproduced the action exactly as they saw it. Smith and Bryson (1998) and White and Brown (1998) reported similar results for some of their gesture imitation tasks.

Relation Between Social-Cognitive Skills and Language

Joint attention and the other social-cognitive skills just discussed all have several things in common. First, to varying degrees, they all rely on children's ability to take the perspective of their social partners toward some object or event, or toward themselves. Second, they require children to attribute to others such mental states as attention, interest, affect, intentions, and prior experience in relation to the object or event at hand. Third, they reflect children's ability, and motivation, to share in those mental states with others. Finally, in typically developing children, at least, the skills are theoretically (and empirically, in the case of joint attention) demonstrated to be essential to typical language development. Deficits in these skills thus would be expected to result in a wide range of problems for language acquisition (see Chapter 5 for further discussion).

That children with autism have difficulties with both language and joint attentional skills is already suggestive of a link between social-cognitive skills and language, but there is more specific and direct evidence as well. Individual differences in the joint attentional skills of children with autism correlate with individual differences in language skills. This would not have to be the case
because it easily could be that a certain level of joint attentional skill is needed for basic language abilities, but then further levels of joint attentional skill result in no further increases in language abilities. It is also the case that skills of other types are necessary for typical language abilities (e.g., skills of concept formation and categorization), and these could easily mitigate quantitative correlations between joint attention skills and language development. Still, a number of studies have shown that these two key skills do indeed correlate strongly.

Robust, positive correlations (ranging from $r = .43$ to $r = .94$) consistently are found between various measures of joint attention and various measures of children's language. Several researchers (e.g., Landry & Loveland, 1988; Mundy et al., 1990; Mundy, Sigman, Ungerer, & Sherman, 1987; Wetherby et al., 1998; see also Chapter 4) have found correlations between children's overall joint attention scores (and individual components of those scores) and measures of concurrent and later receptive and expressive vocabulary sizes. It is interesting to note that Mundy and colleagues (1994) also found that children's joint attention scores were negatively correlated with parents' reports of disturbances (including odd prosody and echolalia) in their child's language. Furthermore, the relation between joint attention and language extends beyond vocabulary size, as strong, positive relations also are found between

1. joint attention and acquisition of grammar (Rollins & Snow, in press, found that 89% of the variance in children's syntactic development was explained by joint attention scores)
2. joint attention and correct use of Italian pronouns (Landry & Loveland, 1988; Loveland & Landry, 1986).

These joint attention-language correlations are specific to joint attention and are not a reflection of general social ability: Few relations are found in any of these studies between general social interaction or communicative behavior regulation (i.e., requesting) measures and language.

In addition to correlations between joint attention and language, several studies have found correlations between imitation, another important social-cognitive skill, and language development for children with autism. These children's vocal and gestural imitative skills are positively related to their verbal and nonverbal communication skills (Abrahamsen & Mitchell, 1990; Curcio, 1978; Dawson & Adams, 1984).

There are thus correlational relations between various social-cognitive skills and language development in children with autism. Most of the relations are with children's vocabulary size, a somewhat blunt measure of children's communicative competence, but there is evidence that these social-cognitive skills also are important in other, very specific areas of language acquisition and use by children with autism. For example, another parallel between joint attention and language concerns the function of joint attention and language. Children with autism tend to talk less often for the purpose of sharing or seeking information than for the purpose of expressing needs and wants (Stone & Caro-Martinez, 1990; Tager-Flusberg, 1992, 1993; Wetherby et al., 1998; Wetherby & Prutting, 1984). This is not surprising given that 1) the function of joint attention is to share something—affect, interest, and attention—about objects with others and 2) children with autism engage more for imperative purposes than declarative ones.

Another interesting characteristic of the language of children with autism that reflects joint attention deficits is these children's tendency occasionally to learn words for incorrect referents. Baron-Cohen, Baldwin, and Crowson (1997), in a replication of Baldwin's (1993a, 1993b) discrepant labeling studies, found that children with autism did not use the adult's gaze direction to learn new words, instead learning the new word for the object they themselves were looking at when the word was uttered. Because of the tendency of these children not to use joint attention behaviors to establish reference and because adults occasionally label objects that are not already in the child's focus of attention, it is not surprising that there are abundant examples of "metaphorical speech" in children with autism (Baron-Cohen et al., 1997; Kanner, 1973). These types of errors can be avoided if caregivers follow into children's focus of attention when using new words: Watson (1998) found that language comprehension of children with autism was positively correlated with frequency of maternal follow-in language and negatively correlated with maternal direct-speech language (as is the case with typically developing children). Thus, whereas these children may learn words in scaffolded interactions in which adults follow into their focus of attention, they are less able to do this in less ostensive contexts, when they must determine and use the speaker's perspective.

Therefore, children's ability or tendency to engage in joint attention is related to their language skills in many ways. A similar pattern is found for deficits in understanding others' communicative intentions. For children who do not understand communicative intentions, in the extreme case, language would simply be semirandom sounds that came out of others' mouths (the sounds would not be completely random because sometimes they might be associated with certain objects or situations). If such children were to speak at all, their speech would be limited to a small number of familiar nouns or phrases, learned associatively or through training and reinforcement, which more often than not would involve highly motivating, imperative situations (e.g., "I want juice"); this type of speech is characteristic of a subset of children with autism. If children understood a little more about others' intentions with regard to language—that is, others' informative intention but not their communicative intention (see, e.g., Happé, 1993)—then language would have more meaning but nevertheless would be incomprehensible in some very specific circumstances. For example, such children would not understand that there can be another level to language—the level of communicative intention—and this would be manifest by their tendency to mistakenly interpret all language literally. Thus, these children should show specific difficulties with
comprehension of such figurative aspects of language as metaphors, irony, sarcasm, jokes, and lies—but should not have difficulties with similes or other literal language, because in the latter cases, understanding the added layer of communicative intention is not necessary. There is abundant evidence of these specific strengths and weaknesses in the comprehension of children with autism (e.g., Happé, 1993; see also Frith, 1989, for further discussion). Similar results are found with production of language as well: Eales (1993) has demonstrated that although the inappropriateness of execution of informative intentions of adults with autism was no different from that of adults with receptive language disorder, adults with autism produced significantly more utterances reflecting failures of communicative intentions and relevance than did the other adults.

A similar pattern also is found for role-reversal imitation. The tendency of children with autism to engage in echolalia, or direct imitation of others' speech, is well known, and we would argue that this tendency is a reflection of these children's ability to imitate but not to engage in role-reversal imitation—these children learn words and phrases exactly as they hear them. Echolalia may be used communicatively by children with autism (Prizant & Duchan, 1981), but it clearly does not involve flexible, reciprocal understanding of language. Echolalia decreases as children's language level increases (McEvoy, Loveland, & Landry, 1988).

In many cases, early speech (i.e., in the one- and two-word stages) learned via simple imitation might be indistinguishable from speech learned via role-reversal imitation. Once children begin including personal pronouns and other deictic words (i.e., words whose meaning changes depending on context; e.g., *you*, *this/that*) in their speech, however, the type of imitation used to learn the words becomes more obvious. If children then use the same form the speaker did in these cases, saying, for example, “You want candy,” to ask for candy, this provides a clear indication of imitation without speaker–listener role reversal. Children with autism often make these types of errors. Furthermore, there is experimental evidence that children with autism can comprehend *me/you* pronouns in others' speech but have difficulty then using those pronouns in their own speech (i.e., reversing the listener's and speaker's roles). Jordan (1989) demonstrated that children with autism responded correctly to instructions such as “Put the hat on you/me” but did not usually respond with the correct pronoun when completing the speaker's statement, e.g., “Look! The dog has jumped on...?” Most of children's “errors” in this case were caused by substituting the experimenter's or their own name (appropriately; e.g., if the dog jumped on the child, the child used his or her name instead of the experimenter's name) for the pronoun, a strategy that does not require role reversal of adults' language.

We have emphasized so far the importance of several early social-cognitive skills for language acquisition. Related to each of the social-cognitive deficits we have discussed is the well-known deficit in theory of mind (ToM) exhibited by children with autism (Baron-Cohen, Tager-Flusberg, & Cohen, 1993). Several researchers (e.g., Baron-Cohen, 1988; Frith, 1989; Tager-Flusberg, 1993, 1997) have described theoretical and empirical relations between ToM abilities and various aspects of (more advanced) language use. For example, understanding of others' thoughts, beliefs, knowledge, interests, emotions, and other mental states is particularly important in the area of pragmatics. Children with autism have been shown to have deficits in a wide range of pragmatic skills: For example, they do not take into account what is new, interesting, or important to the listener; they ask socially inappropriate and embarrassing questions; and they use pedantic speech. Both the narratives of children with autism (see Loveland & Tuni, 1993, for a review) and the instructions the instructions they give to others (Loveland, Tuni, McEvoy, & Kelley, 1989) also are inadequate, in general, for the same sort of reason: These children do not take into account what listeners know or do not know. Indeed, direct relations have been found between ToM and such aspects of language as understanding of figurative language (Happé, 1993), narrative skills (Tager-Flusberg & Sullivan, 1995), and various pragmatic skills (Eisenmajer & Prior, 1991).

There are several other characteristics of the language of children with autism that may result from difficulties with ToM, perspective taking, and keeping track of what the listener knows. For example, children with autism use fewer words for mental states than do other children (e.g., Tager-Flusberg, 1992). Furthermore, although the form of the syntax of children with autism is in general unimpaired (Cantwell, Baker, & Rutter, 1978a; Tager-Flusberg et al., 1990), their use of various syntactic forms differs markedly from that of other children in specific ways that reflect difficulties understanding that others may have different knowledge and points of view from their own (see Tager-Flusberg, 1997, for a review). Children with autism also often do not choose pragmatically loaded words (e.g., *all*, *this/that*) and forms (e.g., past tense) correctly (Bartolucci, Pierce, & Streiner, 1980), and they also do not stress important parts of sentences communicatively (e.g., Fine, Bartolucci, Ginsberg, & Szatmari, 1991).

CLINICAL AND EDUCATIONAL IMPLICATIONS

The joint attention deficits in autism are well known and are already often a target of intervention programs. In addition to promoting joint attentional skills per se, therapists should work with caregivers on establishing routine interactions (see Weithby, 1986, and Chapter 9 for further discussion) and on
following into their child's focus of attention when using new language (or at least work on not directing children's attention to something new as often; Watson, 1998) because children with autism undoubtedly need both kinds of scaffolding even more than typical beginning language learners do. Following into the focus of attention of these children is especially important because they do not respond to others' initiations well—when adults attempt to direct their attention to new objects, they usually do not follow (McArthur & Adamson, 1996; Leekam et al., 1997). Although it is possible to draw these children's attention to a new object by using literal instead of conventional methods (McArthur & Adamson, 1996), once this is accomplished, it is important to encourage the sharing of attention with other people, perhaps through the use of the kinds of tricks that parents of typical infants use when joint attention is first developing (e.g., bringing the object into the joint line of regard of the child and the adult; Bruner, 1983).

In addition to joint attention, we believe that other factors are important in language acquisition: understanding communicative intentions and role-reversal imitation. These skills should also be targeted for intervention. For example, work on understanding others' communicative intentions could begin with work on understanding others' behavioral intentions (e.g., encouraging children to figure out why others behave the way they do, calling attention to non-objects and mistakes). Work on imitation should be extended beyond having children mimic simple actions or vocalizations without perspective taking. An important accomplishment to traditional imitation tasks is the addition of tasks that require some form of role reversal, such as a hiding/game in which the adult plays one role and then at a later time is asked to play the game in the other role.

Finally, it is important for clinicians and parents to think about language and communication in a more cultural, social-pragmatic way, as opposed to teaching and rewarding associations between words and objects (see Seibert & Oller, 1981; Wetherby, 1983; and Chapter 9). Only with such an approach will children be able to learn language flexibly and pragmatically.

DIRECTIONS FOR FUTURE RESEARCH

There is a lot that is not yet known about the important abilities discussed in this chapter. More information is needed about all of these abilities in children with autism and in other children. For joint attention, it is necessary to know why children with autism do not tend to engage in joint attention behaviors. Various theories have been proposed (see, e.g., Coughlin et al., 1994; see also Chapter 7) but little testing of the underlying mechanisms of joint attention has been conducted. For communicative intentions, tests of the prediction that even highly verbal children with autism would have difficulties in the different kinds of word-learning studies conducted by Baldwin (1993a, 1993b) and Tomasello and colleagues (e.g., Akhtar & Tomasello, 1996; Tomasello & Barton, 1994; see also Tomasello, Call, & Gluckman, 1997) would be valuable. For role-reversal imitation, direct tests of this ability are needed, both for gestural/object imitation and for imitation of language. It also would be interesting to see whether the same relations between deficits in these skills and specific weaknesses in language are apparent in other populations, such as blind children.

CONCLUSIONS

Several reviews of language in autism (e.g., Baron-Cohen, 1988; Frith, 1989; Tager-Flusberg, 1993) have proposed that the specific pattern of strengths and weaknesses shown by people with autism in their language development fits well with a ToM explanation. We agree that having a ToM is necessary to produce appropriate narratives, to understand figurative language, and to engage in relevant, interesting conversation with others. We also point out, however, that earlier in the process of language acquisition, children are just beginning to learn words and other linguistic symbols, deficits in other, more foundational social-cognitive skills—specifically, joint attention, understanding of others' communicative intentions, and role-reversal imitation—can explain many of the better selective impairments in the language of children with autism and that these skills, especially understanding others' communicative intentions, still are important in later language as well.

REFERENCES


Joint Attention and Cultural Learning


4

Joint Attention, Social Orienting, and Nonverbal Communication in Autism

Peter Mundy and Jennifer Stella

Autism is a biologically based disorder that may be more prevalent than once thought, occurring at a rate of 1 in 1,000 (Bryson, 1996). It is characterized by impaired social development (Bailey, Philips, & Rutter, 1996; Kanner, 1943). Rather than display a “pervasive lack of responsiveness to others” (American Psychiatric Association, 1980), however, people with autism are now understood to display a pattern of strengths and weaknesses in the acquisition of social and communication skills, which changes with development (Mundy & Sigman, 1989a). Understanding the nature of this pattern of strengths and weaknesses may be fundamental to research and intervention with these children.

In preverbal children, the communication disturbance of autism is exemplified by a failure to adequately develop joint attention skills. These skills involve the tendency to use eye contact, affect, and gestures for the singularly social purpose of sharing experiences with others. Prototypical of joint attention behavior is the act of pointing or showing to share one’s pleasure in a toy. Alternatively, less impaired is the use of eye contact and gestures to regulate the behavior of others for more instrumental purposes. These behaviors include requesting aid in obtaining objects or even displaying attachment/reunification behaviors after a caregiver separation (Curcio, 1978; Loveland & Landry, 1986; Mundy, Sigman, Ungerer, & Sherman, 1986; Sigman & Mundy, 1989; Wetherby, Prizant, & Hutchinson, 1998).

Portions of this chapter have been reprinted from “On the nature of communication and language impairment in autism,” Mundy, P., & Marcus, J., Mental Retardation and Developmental Disabilities Research Reviews, 3, 343-349, Copyright © 1997, Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc., adapted by permission.

Preparation of this manuscript was supported in part by National Institute on Deafness and Other Communication Disorders Grant No. 09484 and by a grant from the Florida State Department of Education to the University of Miami Center for Autism and Related Disabilities.
Also in the Communication and Language Intervention Series:

Volume 5
Language Intervention:
Preschool Through the Elementary Years
edited by Marc E. Fey, Ph.D.
Jennifer Windsor, Ph.D.
and Steven F. Warren, Ph.D.

Volume 6
Assessment of
Communication and Language
edited by Kevin N. Cole, Ph.D.
Philip S. Dale, Ph.D.
and Donna J. Thal, Ph.D.

Volume 7
Transitions in
Prelinguistic Communication
edited by Amy M. Wetherby, Ph.D.
Steven F. Warren, Ph.D.
and Joe Reichle, Ph.D.

Volume 8
Exploring the
Speech-Language Connection
edited by Rhea Paul, Ph.D.

Volume 9
Autism Spectrum Disorders
A Transactional Developmental Perspective
edited by
Amy M. Wetherby, Ph.D., CCC-SLP
Professor
Department of Communication Disorders
Regional Rehabilitation Center
Florida State University
Tallahassee

and
Barry M. Prizant, Ph.D., CCC-SLP
Director
Childhood Communication Services
Cranston, Rhode Island
Adjunct Professor
Center for the Study of Human Development
Brown University
Providence, Rhode Island
Contents

Series Preface ............................................................... vii
Editorial Advisory Board ............................................... ix
Contributors ............................................................... xi
Acknowledgments .......................................................... xvii

1 Introduction to Autism Spectrum Disorders
   Amy M. Wetherby and Barry M. Prizant .......................... 1

Part I Theoretical and Research Foundations:
Understanding the Developmental Context of Autism

2 Diagnosis of Autism Spectrum Disorders in Young Children
   Catherine Lord and Susan Risi ................................... 11

3 Joint Attention, Cultural Learning, and Language
   Acquisition: Implications for Children with Autism
   Malinda Carpenter and Michael Tomasello ................... 31

4 Joint Attention, Social Orienting, and Communication
   in Autism
   Peter Mundy and Jennifer Stalla ................................. 55

5 Intersubjectivity in Autism: The Roles of Imitation and
   Executive Function
   Sally J. Rogers and Loisa Bennett.............................. 79

6 Understanding the Nature of Communication and
   Language Impairments
   Amy M. Wetherby, Barry M. Prizant, and
   Adriana L. Schuler .................................................. 109

7 Sensory Processing and Motor Performance in Autism
   Spectrum Disorders
   Marie E. Anzalone and G. Gordon Williamson ................ 143

8 Neurological Underpinnings of Autism
   Natasha Akshoomoff ................................................ 167