14 Learning to communicate

VIKRAM K. JASWAL and ANNE FERNALD

Introduction

The ability to communicate is essential for all animals living in social groups. Some social species such as ants and bees communicate through intricate but relatively inflexible behaviours that change little over the lifespan. At the other end of the spectrum, primates depend on the adaptable use of continuously varying communicative behaviours that change fundamentally between infancy and adulthood. As different as they are, ants, bees, chimps, and humans have all inherited the need and the potential to interact successfully with *conspecifics* (members of the same species), although the behaviours and mental faculties required for communicative competence in these species are obviously very different. In contrast to the fixed behavioural patterns of social insects, the development of communication in primates requires an advanced capacity for learning. Young chimps and humans learn to convey information relevant to their needs through vocal, facial, and motor behaviours, and learn to monitor the behaviours of others in order to read their intentions and make sense of their actions, skills crucial for success in the complex social worlds in which they live. By the end of their first year, human infants move beyond the non-verbal behaviours shared with other primates, toward a medium of communication that is incalculably more powerful and flexible. This chapter focuses on how human infants learn to communicate, at first through voice, face, and gesture, and gradually through the use of language.

Infants are social creatures from the beginning, by virtue of being born into communities concerned for their survival. Caretakers in every culture tend to interpret babies' cries as a sign of discomfort and to respond accordingly, although the cry of the newborn is certainly not an intentional communication. In his theory of attachment (see Chapter 12), John Bowlby drew on insights from animal behaviour to interpret the social capabilities of human newborns from an evolutionary perspective. Bowlby pointed out that the cries and other reflexive behaviours of the young in many species serve an evolutionary function by alerting caregivers to the infant's needs. Such signals are particularly important in species where the young require a prolonged period of parental support. In '*precocial*' species such as horses, sheep, and geese, the young are

born motorically well developed, able to move around on their own, and seek food and protection by following the parent shortly after birth. However, in *altricial* species such as cats and robins, the young remain helpless for a period and need the parent to stay close by to provide food and protection. Although infants in altricial species are too motorically immature to maintain proximity by following the parent, species-specific social signals enable them to achieve that goal by other means. The gaping mouth of the herring gull chick elicits feeding behaviour from the parent, and the ultrasonic distress cry of the rat pup inspires the mother to return to the nest and lick the pup, both examples of the complementary evolution of infant behaviours matched by appropriate parental behaviours. Bowlby extended these ethological observations to humans, the most altricial species of all, reasoning that the earliest social signals are cries, smiles, and other reflexive behaviours indicative of the infant's internal state. These communicative signals have evolved as the involuntary tactics of the helpless newborn for keeping the parent close at hand.

What is communication?

Although biologists, linguists, psychologists, and engineers still debate about how best to define communication (Mellor, 1990), they generally agree that communication occurs when a signal given by one organism is perceived by and influences the behaviour of another organism. But what exactly is a signal? When the uneven gait of an injured zebra reveals vulnerability to a hungry lion, is this a social signal? Signals convey information, and the zebra's behaviour is certainly informative. However, most ethologists (scientists interested in natural behaviour) would agree that the wounded animal's gait is a direct consequence of a physical condition, not a signal specialized for communication. What about the pheromone released by a sexually receptive female cat that attracts all the male cats in the neighbourhood? Some would consider this a social signal because it reflects an evolved mechanism for broadcasting important information to conspecifics, although others might object that it too is a by-product of the female's physiological condition. The alarm call of a frantic monkey on seeing a predatory snake also conveys information about the monkey's state of arousal; however, this behaviour more clearly qualifies as a social signal because it has evolved to convey information to conspecifics and because it is turned on and off in particular circumstances (Hauser, 1996). How do the cries and smiles of the human infant compare with these examples? In the first months of life, cries and smiles are behaviours that are closely linked to the infant's physiological states and can be perceived and interpreted reliably by other humans. Thus they are informative and can strongly influence the behaviour of conspecifics, as in the example of the female cat. However, unlike the monkey's alarm calls, these early cries and

smiles do not yet relate to anything in the outside world beyond the infant's immediate sensations.

Although in terms of their communicative competence human newborns may start out roughly at the level of a cat in heat, they begin to overtake the monkeys within the first months of life. The monkey's alarm call seems different from the behaviours of the zebra and the cat, somehow closer to human communication, but why? Several questions come to mind: Did the monkey realize the snake might attack? Was the alarm call a deliberate warning to others in the troop informing them of potential danger? Did the monkey understand that the alarm call would cause the other animals to seek safety? Did the other monkeys understand that the caller had seen a snake? These questions get at whether the monkey's communicative signal was intentional, was perceived by others as intentional, and was motivated by assumptions about the intentions and future actions of others. In the case of a human shouting 'Watch out for the snake!' the answer to all these questions would be positive, although with monkeys it is much less clear (Cheney & Seyfarth, 1990). This example makes the point that intentionality and an understanding of the mental states of others are critical in distinguishing human from animal communication.

In the first year of life, human infants move beyond their early reflexive cries and smiles toward an understanding of other minds that enables them not only to interpret the cries and smiles of others, but also to begin to communicate through language. Although the awareness of other points of view and the ability to make sense of meanings through spoken language emerge only gradually over this period, the speechless infant is still fully engaged in human communication on other levels. Infants express their sociability in the early months not only through their own emotional expressions, but also by their intense interest in the voices and faces of others. Bowlby referred to this as a universal stage of 'indiscriminating' social awareness, because months before infants develop strong emotional attachments to particular individuals, they appear to be fascinated by people in general. This observation has now been substantiated by 30 years of experimental studies exploring infants' early sensitivities to faces, voices, gestures, and the dynamics of social interaction, to be reviewed briefly in the following sections. Six month olds may not yet understand that other people have emotions that lead them to act in certain ways and not others; however, by this age they are already very attentive to the emotional signals that will ultimately enable them to make such inferences.

Listening to voices

While still *in utero* (in the womb), infants are able to hear the intonation of the mother's voice: in Chapter 3 it was noted that the mother talking will be

readily heard by the fetus, and her voice is an auditory signal potentially rich in information. After hearing a voice for only a few seconds, experienced listeners make rapid judgements about whether the speaker is a friend or a stranger, male or female, young or old, perhaps also concluding that the speaker is angry or tired or ill. Research with newborns shows that even listeners with minimal experience encode socially relevant characteristics of vocal signals. DeCasper and Fifer (1980) asked whether infants can recognize their mother's voice at birth, testing newborns in a conditioning procedure where they learned to adjust their sucking response in different ways to produce either their own mother's voice or the voice of another woman. Infants only a few hours old chose more often to listen to the recording of their own mother. As the newborns had almost no postnatal experience hearing the mother speak, and no opportunity to associate her voice with pleasurable experiences such as nursing, the listening preference for the familiar voice was apparently based on prenatal experience.

Do these results show that newborns can recognize the mother through her voice? Certainly not in the sense that an adult can identify an individual speaking on the phone, nor even in the sense that an 8 month old shows recognition by smiling in response to the mother's voice and looking toward the door in anticipation of her appearance. In these examples, 'recognition' implies that the familiar voice is part of a complex schema based on extensive experience with the individual. For the inexperienced newborn, however, the mother's voice is merely an auditory stimulus that has been experienced before but not yet in association with any other aspects of her identity in postnatal experience. Although this study demonstrated recognition in only a limited sense, the DeCasper and Fifer (1980) findings were exciting because they were the first to show that even before birth, infants are attentive to human voices. These researchers also showed that newborns can distinguish two verses spoken in the same voice, one read aloud by the mother several times during pregnancy and the other a verse never heard before (DeCasper & Spence, 1986). Given this choice, newborns chose to listen to the verse they had heard while *in utero*. As the mother was the speaker in both cases, the two verses could not be discriminated on the basis of acoustic characteristics unique to her voice, but only on the basis of rhythmic and other prosodic differences peculiar to each. Newborns can also distinguish one language from another, preferring to listen to the language they have been hearing prenatally (Mehler *et al.*, 1988; Moon *et al.*, 1993). These findings reveal that newborns are already capable of extracting information from voices along multiple dimensions that will be socially and linguistically relevant in postnatal life.

After birth, the infant begins to experience voices in the context of increasingly rich forms of social stimulation. In many cultures, adults interacting with infants use a special form of speech that is more lively and musical than the speech typical of adult conversation. Figure 14.1 shows the vocal melodies of an American mother speaking to her 4-month-old child in the

infant-directed (ID) speech style (sometimes called *Motherese*), and also speaking to an acquaintance in the adult-directed (AD) speech style. Note how the intonation contours of ID speech are greatly expanded in pitch range. Analyses of ID speech in British and American English, French, German, Italian, Japanese, and Chinese show that ID speech in these languages is typically higher in pitch with more exaggerated intonation contours, shorter utterances, and longer pauses than in AD speech (Griesser & Kuhl, 1988; Fernald *et al.*, 1989). When given the choice of hearing ID or AD speech in an auditory preference experiment, infants listen longer to ID speech (Fernald, 1985) and show more positive emotion (Werker & McLeod, 1990). Through social interaction over the early months, the melodies of the mother's voice become associated with playful interaction, comforting, feeding, and many other pleasurable aspects of parent–infant interaction, which may account in part for the listening preference for ID speech shown by 5 month olds. However, even newborns show this preference (Cooper & Aslin, 1990), suggesting that the vocal melodies of ID speech are appealing in themselves without extensive experience.

Through experience, however, these intonation patterns take on meaning for the infant because parents in different languages tend to use characteristic intonation patterns in particular interactional and emotional contexts

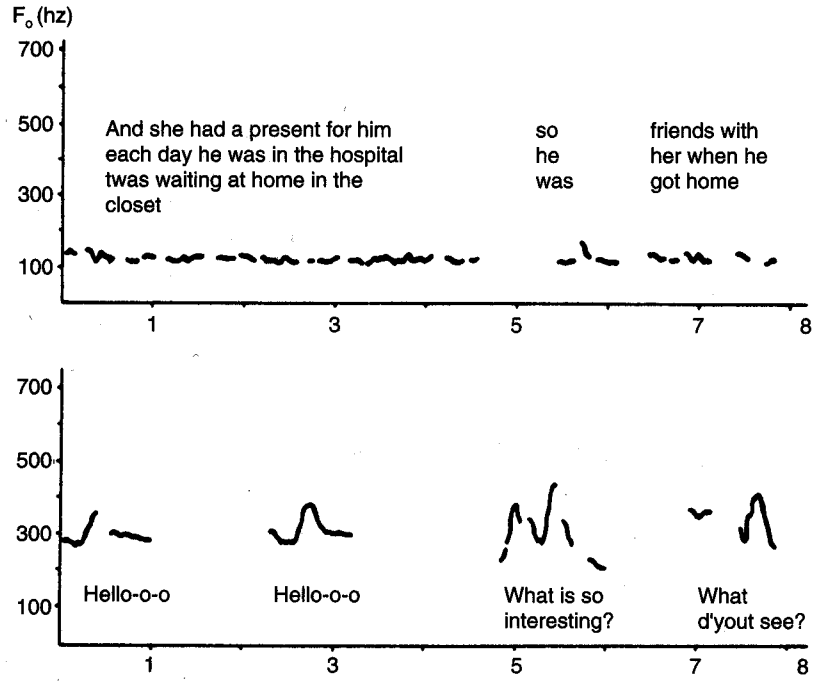

Figure 14.1 Intonation contours from the adult-directed (above) and infant-directed (below) speech of an American mother. Graph shows the movements in fundamental frequency (F_0), which corresponds to the pitch of the mother's voice.

(Fernald, 1992). For example, the intonation contours used to praise an infant are typically wide in pitch range with a rise–fall pattern, quite different from the typical contours used for warning or prohibition, which are shorter, more abrupt, and lower in pitch. Contrast how you might say 'What a good baby!' with how you would say 'No! That's not for you to play with!' By the age of 5 months, infants respond distinctively to these different contour types, even in a language they have never heard before. When English-learning infants heard 'praise' and 'prohibition' contours in German and Italian ID speech, they smiled more in response to the vocalizations that were positive in tone and looked wary in response to the negative vocalizations (Fernald, 1993). As the words were unintelligible in any case and were spoken in an unfamiliar language, these findings show that certain culturally widespread features of ID intonation are effective in eliciting emotion in preverbal infants. However, even though 5 month olds respond appropriately to intonation contours associated with communicative intentions such as praise and prohibition, we cannot conclude that they *understand* anything about the speaker's intentions at this age. In terms of communicative competence, these early selective responses based on intonation show an increasing sensitivity to emotion in the voice, but they are still not fundamentally different from the selective responses of monkeys to alarm calls.

Making sense of faces

Infants find faces as fascinating as voices from the beginning, and facial expressions provide another rich source of social information in human communication. Chapter 10 reviews the questions scientists have asked about how infants first process faces as visual stimuli and then gradually learn to interpret the emotional information available in facial expressions. Are infants biologically predisposed to find human faces interesting, or does the 'specialness' of faces emerge through experience? Given the poor visual acuity of infants in the early months, what kinds of visual information can they extract and remember when looking at faces? When are infants able to categorize and interpret emotional expressions in the face? These questions are all relevant to the development of communication, for an obvious reason: the ability to recognize individual faces and to interpret facial expressions is only useful when another person is present, i.e., in the context of social interaction. Building on the research described in Chapter 10, we consider another important question: How do infants begin to *use* their growing knowledge about faces in the process of communicating with other people?

By the age of 8 months infants can tell that happy faces are different from angry faces in terms of their visual features (Ludemann & Nelson, 1988). While this ability may be necessary for interpreting the emotional meaning of

these expressions, it is far from sufficient. To interpret an expression in the flow of social interaction requires integrating information from multiple sources and making rapid inferences about what just happened and what will happen next. For example, when you see a big, sudden smile on your friend's face as she looks behind you, your response is not to conclude dispassionately 'Oh, the position of Sarah's mouth just changed' and leave it at that. Most likely you will make several automatic and instantaneous attributions, inferring that Sarah's expression indicates an internal state of pleasure, that she has seen something that caused the change in state, that whatever she has seen is something she regards positively, that she is more likely to approach than to flee from whatever it is she is looking at, and so on. As you make these rapid inferences you will probably also turn to follow her gaze so you can share her focus of attention. Without a word spoken and in a fraction of a second, you have processed several kinds of information from your friend's face that enabled you to make assumptions about aspects of the situation that you did not (and could never) experience directly, i.e., how Sarah is feeling, what made her feel that way and why, and what she might do in response. This kind of 'mind-reading' pervades human communication on every level, and by the end of the first year infants are well on their way to developing the complex mental abilities necessary for understanding other minds.

In the example above, at least three kinds of information could be extracted from what was happening on Sarah's face: first, the change in expression, marking the onset of the causal event; second, the expression itself, indicating the nature and intensity of her internal emotional state; and third, the direction of her gaze, indicating attention to something presumably involved in the change of state. Co-ordinating these different sources of information in a moment and initiating an appropriate response is a challenging task that demands much more than the ability to categorize static facial expressions. In their pioneering study of 'social referencing', Campos and Stenberg (1981) were the first to explore how 12-month-old infants use these emerging mind-reading skills in action. Infants were placed on a 'visual cliff', an apparatus originally developed for testing depth perception, which gives the compelling illusion of a sudden drop. In fact the chasm is spanned by a sturdy plexiglass surface to support the infant, but most 12 month olds stop at the brink and refuse to venture further. Positioned on the other side of the gap, the infant's mother was instructed to pose one of two facial expressions—a big smile or an exaggerated fear face—when the infant stopped at the brink and looked across at her. If the mother smiled, most infants overcame their hesitation and proceeded across the chasm; however, if the mother displayed fear, all of the infants retreated from the edge and refused to cross. Campos and Stenberg concluded that in situations of uncertainty, infants check the mother's face to see her appraisal of the situation and then decide how to respond depending on the emotion she is displaying. This compelling phenomenon was dubbed *social referencing* because infants appeared to be referring to a social partner

for non-verbal guidance, modifying their behaviour based on the mother's positive or negative appraisal of a potentially dangerous situation.

Research on social referencing suggested that by the end of their first year, infants can engage in complex reasoning processes, which integrate many sources of non-verbal information to make inferences about the mother's emotional state. On the surface, these conclusions might seem plausible, but there are reasons to prefer a more cautious interpretation. For example, imagine a wary infant who hesitates to approach a fascinating but unfamiliar mechanical bear; the infant then turns to the mother who displays fear as she looks at the same toy. If the child avoids the bear after looking at her face, can we really conclude that this response is guided by inferences about the mother's internal state? We can assume that the mother's fear expression, a kind of visual alarm call, is motivated by more complex processes than the monkey's alarm call because she is not actually afraid of the toy herself. However, the infant's avoidance response may result directly from discomfort at seeing the strange expression, or from past experience with fear faces followed by unpleasant events. These explanations do not presuppose any implicit reasoning about the mother's emotional state, i.e., that if the mother has this particular facial expression as she looks at the bear, she must be afraid of the toy and thus it should be avoided. If simpler explanations can account for the infant's avoidance of the toy, they imply no more (or less) mind-reading ability than monkeys use to flee in response to the alarm call of a conspecific.

We cannot resolve these complex questions here (see Baldwin & Moses, 1996), but the example serves to make some important points. First, while the 3-month-old infant looks to the mother's face purely for the pleasure of interacting with her, the 12 month old may look to the mother's face to seek information about something in the world beyond their interaction. Second, while the response of the 12 month old may seem simple and straightforward, it represents an enormous advance in communicative competence, although the mental processes involved are complex and difficult to tease apart. And finally, our insights into the cognitive and emotional mechanisms underlying the development of communication over the first 2 years do not result from any one study or paradigm; rather, they emerge slowly from many convergent studies of different behaviours, which all reveal the infant's gradual progress in understanding other minds.

Appreciating intentions

Already in the first year of life infants begin to understand that the actions of others are usually purposeful, or goal-directed, and that much of human behaviour is motivated by desires and beliefs (Wellman, 1990). For example, if you were frantically opening and closing the drawers of your desk, rifling

through their contents, your room-mate would likely infer that you wanted to find something, and that you believed it might be in your desk. Even if no words are exchanged, your room-mate will attempt to interpret your behaviour, just as you would attempt to interpret the cause of the big smile on your friend's face described earlier.

Infants could not possibly interpret accurately the reason why you were searching through your desk, as they have no experience with desks, losing things, or searching for them. This is not to say, however, that they lack the drive to interpret the behaviour of other humans. In a series of clever studies by Woodward (1998), infants watched as a hand repeatedly reached toward and grasped a ball in a display containing a ball on one side and a stuffed bear on the other. Recall from Chapters 2 and 5 that infants will lose interest (habituate) when watching the same display repeatedly. Once they had habituated to this action, the locations of the ball and the stuffed bear were switched. If infants paid attention simply to the *path* of the hand as it reached toward an object, they should dishabituate if the hand now reaches toward the new location (which is where the ball now is). However, if infants paid attention to the object or *goal* of the reach, they should dishabituate if the hand reaches toward the stuffed bear, now located at the old location. In fact, this is exactly what infants did: infants as young as 6 months looked longer when the hand reached toward a new goal, but not when it simply reached toward the same goal in a different location. (The same results were obtained when the bear was the first item reached for, and the ball the second.) In other words, 6-month-old infants apparently considered a change in goal to be more noteworthy than a change in path. Even more interesting, when the same study was conducted with a mechanical claw doing the reaching and grasping rather than a human hand, there was rarely a looking preference, but when there was, it tended to favour the new path rather than the new goal. This suggests that by 6 months, infants can distinguish between human-produced behaviour and machine-produced behaviour, and that for behaviours produced by humans, it is the goal that is important.

This early ability to interpret the intentions of other humans in terms of goals is quite striking. However, it is important to caution again against attributing too much to the infant. Even pre-schoolers err on mind-reading tasks that would seem trivial to an adult, often because they can understand behaviour in terms of desires, but have yet to grasp the notion of belief (Wellman, 1990). For example, 3 year olds understand that someone who wants a piece of candy will look for it, but they expect that the person will look for it in the location where it actually is rather than where he or she believes it to be. Thus, if Sally left a chocolate in one location and, unbeknown to her, Anne moved it to a second location, 3 year olds expect that Sally will look for it in the new location—even though she could not have known that Anne moved it (Wimmer & Perner, 1983)! Despite such limitations, an early sensitivity to cues relating to intentionality can help infants as they learn to communicate.

Following gaze and pointing gestures

Prior to about 9 months, infants respond directly to salient sources of stimulation, crying in response to a loud noise or producing a charming smile in response to a game of peekaboo, without attempting to share their experiences with anyone else. At about 9 months, however, things begin to change, as infants try to communicate *about* things to other people For example, while a 6 month old might smile in response to a wind-up toy, a 14 month old would also be likely to look to his mother while smiling, as if to say 'Isn't this cool?' This is an important transition in a child's development, marking the beginning of intentional communication and an understanding of intentional communication by others (Bates, 1979). When older infants vocalize, these signals are often deliberate and may serve no other purpose than to gain the attention of their partner in interaction, quite unlike the cries and babbling of the younger infant. This section focuses on the development of two hallmarks of intentional communication: gaze-following and pointing.

Gaze-following

Following the gaze of another person is a deceptively simple activity. When you stare at an object, it is trivial for another adult to figure out what you have focused your attention on. Once both of you are looking at the same object and thus have established 'joint reference', you can communicate about that object, talking about its features, whether each of you likes it, and so on. Before an infant reaches 8 months, the adult is the one responsible for following the gaze of the infant in order to establish joint reference. That is, when the infant chooses to focus on something, the adult then follows in, frequently labelling and/or commenting on the object of the infant's attention (Collis & Schaffer, 1975). If the adult looks at something different and tries to call attention to it, young infants will frequently look only at the adult's face, apparently unaware of the object of reference.

By about 12 months, however, infants have learned to follow an adult's gaze under most circumstances; if an adult across from the infant looks to the right or left, the infant will generally look in the appropriate direction (Butterworth & Grover, 1988). Interestingly, by 12 months, infants will also turn to follow the look of a decidedly non-human object, provided that the object provides cues to indicate that its actions might be intentional. To demonstrate this, Johnson *et al.* (1998) showed infants a round, fuzzy object that turned either to the left or to the right. The question was whether infants would follow the 'gaze' of this strange object. When the object had what could best be described as eyes, like the object on the right in Figure 14.2, infants did follow its gaze, turning to look in the same direction. Even when the object did not have eyes, like the object on the left in Figure 14.2, infants mirrored its turns, so long as

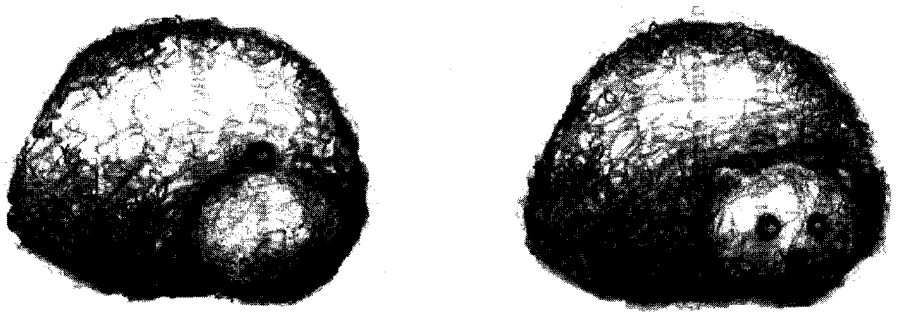

Figure 14.2 Schematic drawing of the novel object with (right) and without (left) eyes. Infants always followed the gaze of the object with eyes, but only followed the gaze of the object without eyes if it had responded contingently to their behaviour. From Johnson *et al.*, 1998; artwork by Kirsten O'Hearn.

it had earlier beeped in response to their vocalizations and flashed attached lights in response to their movements. (Imagine the robot R2D2 from 'Star Wars.') However, in another condition where the object lacked eyes *and* failed to respond contingently, the babies did not follow its 'gaze.' Thus, what drives infants to follow another's gaze seems to be cues related to that individual's intentionality—in this case, eyes and/or contingent responsiveness. By 12 months infants have learned that things with certain cues (i.e., ones human adults construe as intentional) are more likely to provide meaningful information about things in the environment than things without such cues.

Chimpanzees, our closest non-human relatives, can also follow the gaze of others (Povinelli & Giambrone, 2001), suggesting that they may be sensitive to the same cues signalling intentionality as human infants. However, we must bear in mind the caveat from the social referencing work described earlier: infants and chimps alike may use intentional cues to attend to potentially meaningful things in the environment without imputing anything like intentionality to the person (or object) emitting those cues. Eventually, humans do use intentionality to make inferences (e.g., Why is she smiling? Why is he looking in the wrong place?), but the picture is much less clear with chimpanzees (see Premack & Woodruff, 1978; Heyes, 1998). In fact, in experimental tests, chimps do not seem to understand the link between visual perception and knowledge at all: for example, they are as likely to request food from a human trainer wearing a blindfold over her eyes as a trainer wearing a blindfold over her mouth, as if they fail to infer that food would only come from the individual who could see the request (Reaux *et al.*, 1999). Some scientists have argued that the ability to understand another's behaviour as intentional—whether it be through eye gaze or some other action—is, in fact, what enabled human communication as we know it to emerge (Tomasello, 1999).

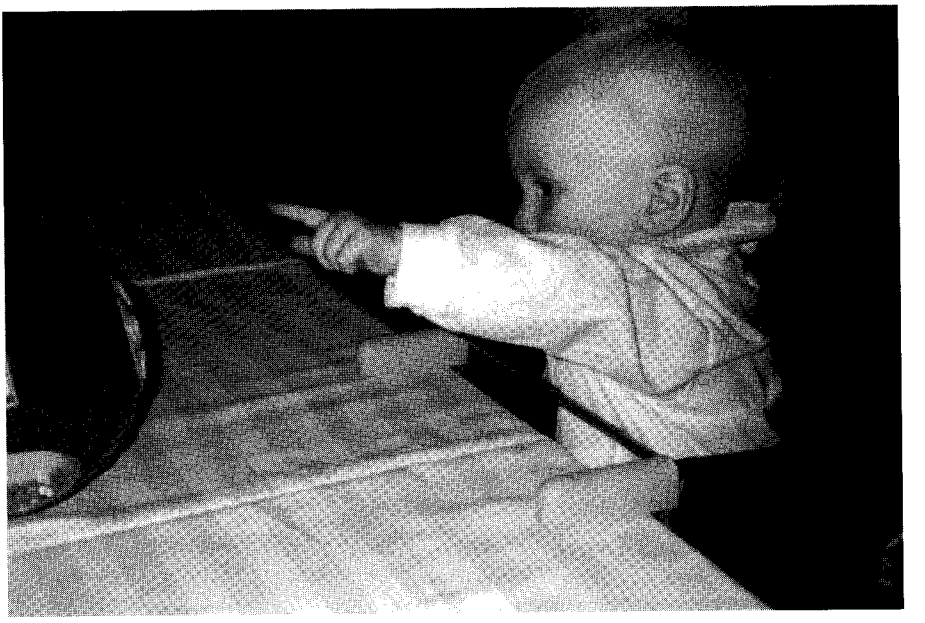

Figure 14.3 Infants begin to point as early as 9 months. Photo by Karen Thorpe.

Pointing

Pointing gestures are another cue that adults frequently provide when directing a child's attention. As with gaze-following, the ability to interpret a point develops early in the second year of life. Prior to about 12 months, infants tend to look at the pointer's hand rather than at the object being pointed to (Butterworth & Grover, 1988). As with gaze, interpreting a point is simpler in some situations than in others: For example, 9 month olds may be able to follow a point as long as the target object is close to the end of the finger, there are no intervening objects, and so on. However, infants are much more flexible by 14 months and can generally follow most pointing gestures successfully (Murphy & Messer, 1977). It is interesting to note that even for adults, pointing in a cluttered environment may not completely disambiguate an object of reference; however, pointing can be helpful in combination with words, which 14 month olds are beginning to understand (Schaffer, 1984).

Some infants begin to use pointing gestures themselves as early as 9 months (Figure 14.3), and most are regularly pointing by 14 months (Schaffer, 1984). Younger infants may point at things even when they are on their own and no social partner is present, suggesting that the earliest form of this behaviour is actually non-communicative (Bates, 1979). However, pointing clearly serves a communicative function a few months later, when the infant not only points but also looks to check that the adult is attending to the point. Bates (1979)

has shown that early communicative pointing, where the infant checks with the adult for confirmation, correlates with a number of language measures of both comprehension and production. Based on these correlations, Bates argues that pointing is really a gestural form of naming and thus is closely related to the development of language skills.

Learning to communicate through language

First sounds

Long before infants produce recognizable words, they spend a good deal of time vocalizing, or babbling—an activity that provides them with valuable practise using the vocal modality. Indeed, the consonants that make up the first 50 words spoken by a child are typically created from sounds that occur frequently in babbling, suggesting that babbled sounds may provide the foundation for the production of a child's first words (Vihman *et al.*, 1985; Vihman, 1992). The development of babbling unfolds in five stages (Oller, 1980; Locke, 1993; Stoel-Gammon, 1998), with transitions from one stage to another depending on the development of neural mechanisms and the vocal tract as well as the practise the child receives in each stage.

In **stage 1** (0–2 months), infants produce reflexive vocalizations, including crying, sneezing, and burping—sounds directly related to their physical state. These vocalizations account for about 90% of all vocalizations at 2 weeks, dropping to 50% by 8 weeks (Stark *et al.*, 1993). In **stage 2** (2–3 months), infants begin 'cooing', so-called because 'coo' is one of the few sounds infants first attempt to produce. In **stage 3** (4–6 months), they start to experiment with the wide array of possible noises they can make with their vocal tract, resulting in a variety of vowel-like sounds as well as hoots and squeals. In **stage 4** (6–7 months), babbling begins in earnest with the appearance of what are called *canonical syllables*. The term 'canonical' can mean 'true to life' and canonical syllables are the first babbling sounds that sound like real words. These consist of a restricted set of alternating consonants and vowels, such as the stereotypical 'gaga,' or 'mama' and 'dada.' Although 'mama' and 'dada' sound like the conventional names for parents in English (and many proud parents treat them as such), we would not call them true words until infants demonstrate an understanding of the correspondence between the sound and its meaning. This is an important caveat to keep in mind, and is related to our earlier discussion emphasizing intentionality as criteria for human communication. At 11–12 months reduplicated babbling ('gaga') gives way to variegated babbling, which involves stringing together different syllables ('bagoo'). Finally in **stage 5** (12 months), infants begin a period of 'jargon' babbling, producing longer strings of consonant–vowel combinations with differing in-

tonation and stress patterns. By this age the rate of babbling is already comparable with the rate of adult conversational speech (Kent & Bauer, 1985). Infants typically produce their first words at about 12 months, just when they are also beginning to use jargon babbling. For several months, babbling and conventional words occur together, but beginning at about 19 months, as children's facility with conventional language increases, the amount of babbling drops off (Stoel-Gammon, 1998).

The restricted set of syllables in early babbling is fairly consistent across linguistic environments and is made up of sounds that occur frequently in the majority of languages (Locke, 1993). This suggests that the range of babbled sounds may be limited more by anatomical or physiological constraints than by environmental factors. However, the ambient language does influence the relative rate at which various sounds appear in babbling (de Boysson-Bardies *et al.*, 1992). For example, babies raised in languages that make frequent use of final consonants (e.g., English) tend to use final consonants in their babbling more than those raised in other languages (e.g., Japanese).

Deaf children also go through a period of vocal babbling, although they seem to be slightly delayed in the onset of reduplicated babbling, and to have a reduced repertoire of sounds. Deaf babies also go through a period of manual babbling, which shares many of the characteristics at the same ages as the vocal babbling of hearing infants (Petitto & Marentette, 1991). In particular, the manual babbling of deaf (but not hearing) infants includes a restricted set of manual babbling gestures, reduplication, and even continuity between the forms used in manual babbling and the first symbolic signs.

First gestures

In addition to early vocal communication, hearing infants also communicate through a host of gestures such as showing, giving, pointing, and ritual requests before they begin to speak their first words (Bates, 1979). These gestures are non-symbolic because they are not used to 'stand for' something else. Despite this, they can still be intentional and communicative if the infant demonstrates an awareness of the effect a particular gesture could have on a conversational partner and persists until the effect is obtained (Bates, 1979). For example, if a 12 month old reaches up repeatedly for a favourite toy high on a bookshelf as her father watches, she might demonstrate intentional communication by looking back and forth between the father and the toy on the shelf, or by vocalizing until the desired goal is achieved. Many of infants' gestures meet these criteria by the age of 10 months. Even after infants begin producing words, they continue to support their linguistic communication with gestures, so gestural schemes are not simply replaced by verbal ones.

Just as babbling may be related to language development in that it provides practise with the communicative modality, regular production of some gestures may be related to language development in that it provides practise

communicating via conventional signals. For example, as mentioned earlier, communicative pointing seems to be correlated with word comprehension at the beginning of the second year, perhaps because both pointing and word comprehension involve establishing joint reference to an external object (Bates, 1979).

Infants may also be able to communicate via symbolic gestures, or gestures that actually do stand for things. In fact, some researchers have suggested that infants can learn to produce symbolic gestures before producing words (Meier & Newport, 1990). One study showed that hearing infants exposed to American Sign Language (ASL) began producing signs at about 8.6 months, 3 months before most children begin producing words (Bonvillian *et al.*, 1983). The gestural modality may show this advantage for a number of reasons, including that motor control of the hands may develop more quickly than control of the articulatory system; signs might be more recognizable by parents than underarticulated words; and many early signs have a high degree of iconicity (Acredolo & Goodwyn, 1990). By iconicity is meant that the sign might look like what it refers to, so that an early sign for 'I'm hungry' might be bringing the hand to the mouth. Before accepting a sign as a symbol, however, we must apply the same standards we require of an early word: it must be intentional and relatively context-independent. When these same standards are applied to the same group of children, the first symbolic signs and the first words both emerge around the age of 12 or 13 months.

For infants who are experiencing difficulty with vocal communication, Acredolo and her colleagues have argued that training them in sign may provide them with a useful communicative outlet (Acredolo & Goodwyn, 1990; Goodwyn & Acredolo, 1993). In one study, they trained mothers to make daily use picture books and other materials in order to demonstrate target gestures (e.g., lip-smacking for 'fish', arm-flapping for 'bird', etc.) to their pre-verbal 11 month olds. Biweekly interviews were conducted to collect information about the children's production and comprehension of these signs as well as the conventional words. On average, infants produced both gestures and words at about the same time, although there was substantial variability among children. In short, some children will progress faster with sign, some with speech, and some will produce words and signs at the same time.

Early comprehension

As you know from Chapter 9, infants have learned an enormous amount about the sounds of their language during the first year. By 8 months infants are attuned to the phonological (sound) system of the ambient or surrounding language (Kuhl, *et al.* 1992; Polka & Werker, 1994), and can recognize recurrent patterns in sequences of speech sounds (Jusczyk & Aslin, 1995). These skills are all essential for recognizing words in fluent speech, which must be

processed at a very rapid rate. However, identifying a sequence of syllables as familiar is just the first step in word recognition, which also requires learning an association between the sound sequence and a particular meaning. Moreover, this sound–meaning association must be more than just a conditioned response to a verbal routine to count as true comprehension.

The question of when comprehension begins has been debated since the earliest scholarly studies of language development 200 years ago. In a diary study published in 1877, Hippolyte Taine noted that when his 11-month-old daughter heard 'Where is Mama?', she always turned toward her mother. This example echoed an observation by Dietrich Tiedemann (1927/1787) published 100 years earlier, whose baby diary was referred to in Chapter 1; when asked to 'Swat the fly', his 8-month-old son also made appropriate gestures, although a different interpretation was offered. While Tiedemann claimed that such responses showed how his son had 'learned to comprehend' simple sentences, Taine was more conservative, suggesting that 'there is nothing more in this than an association'. When his daughter was 12 months old, however, Taine was convinced she demonstrated true comprehension of the word *bébé*. Although her understanding did not appear to coincide with the conventional meaning, Taine claimed that *bébé* nevertheless had 'a general signification' for her which went well beyond a limited association between a sound pattern and a gestural response. These informal observations suggesting that infants begin to demonstrate understanding of words by the end of the first year have been verified by numerous studies in many different languages (e.g., Bloom, 1973; Casselli *et al.*, 1995). One way of gathering data on early comprehension is to ask parents to fill out standardized checklists to keep track of which words their child appears to understand in daily interactions in the home. Figure 14.4 shows that, on one such checklist, infants understand an average of about 50 words by their first birthday, and over 150 by the age of 16 months (Fenson, *et al.*, 1994). As the figure shows, however, there is considerable variability among children on these measures.

While checklists give an estimate of changes in the size and extent of the child's receptive vocabulary over time, they don't reveal anything about important developments in the child's ability to recognize and understand familiar words in continuous speech. If a mother indicates on the checklist that her infant understands the word 'ball' at 12 months, she is likely to check the same box again at 15 months, although in the intervening months there have been dramatic changes in speech processing efficiency that are not obvious in the child's spontaneous behaviour. You are probably not aware of this, but as a fluent language user you typically process 15–30 different speech sounds every *second* in following a casual conversation! How do infants develop the skill to understand spoken language with such remarkable speed and efficiency?

The early development of competence in word recognition has been studied by tracking infants' eye movements as they listen to spoken sentences

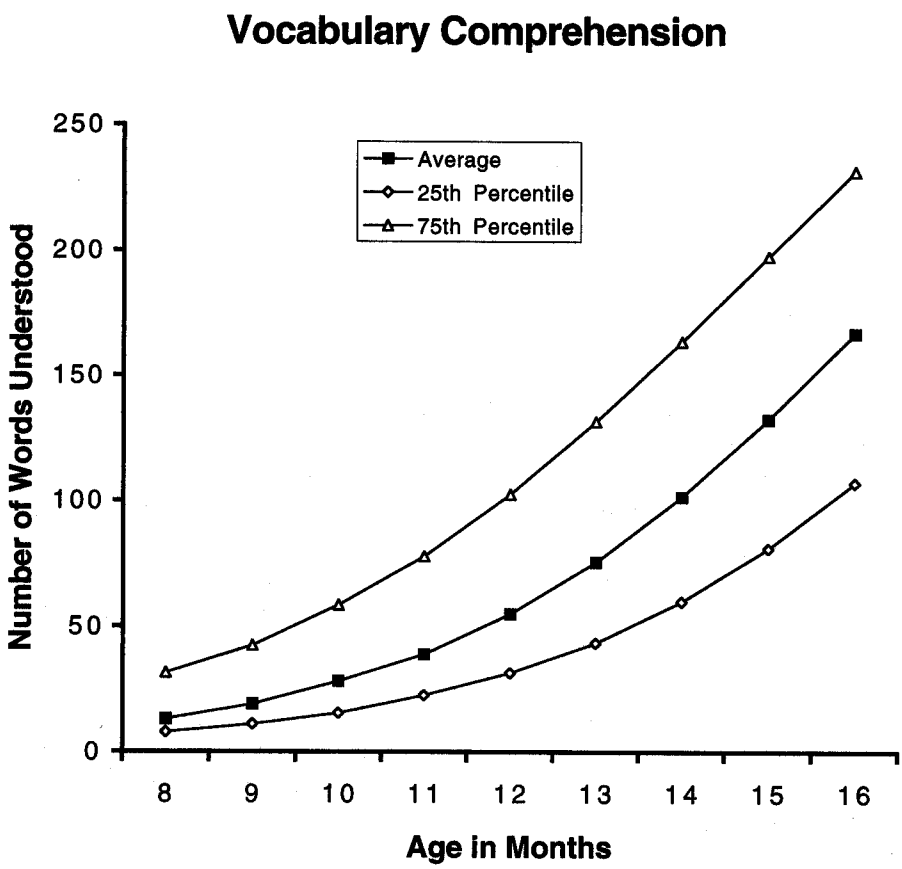

Figure 14.4 Receptive vocabulary size from 8 to 16 months. This figure shows the average number of words infants at various ages understand, and also includes the 25th and 75th percentile range in order to show the large amount of variability in early vocabulary size. Adapted from Fenson *et al.*, 1994.

containing familiar words. Sitting on the parent's lap, infants are observed in a booth where they look at pictures while hearing speech naming one of the pictures, as shown in Figure 14.5. The infant's eye movements are video-recorded during the test session, and then analysed later in slow motion, frame-by-frame, to measure very precisely how quickly the infant's eyes moved to the named picture. By tracking infants' gaze patterns in the process of understanding, it is possible to document impressive progress in the efficiency of spoken word recognition occurring over the second year of life (Fernald *et al.*, 1998). Figure 14.6 shows the average speed of response, or reaction time, for infants at 15, 18, and 24 months of age, as they orient to the appropriate picture in response to hearing the name of the picture. The bar graph in this figure is aligned with the waveform from one of the stimulus sentences, showing the time course of infants' responses in relation to the famil-

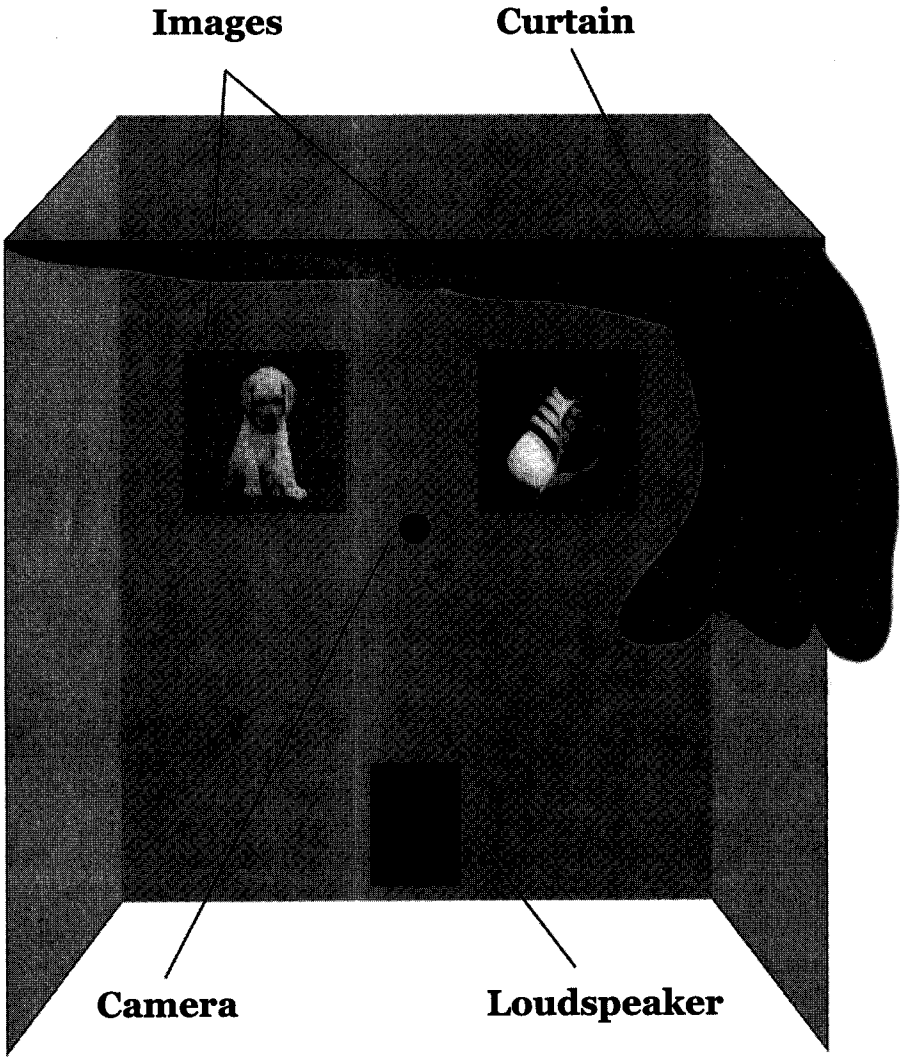

Figure 14.5 Experimental set-up used by Fernald *et al.* (1998). Infants listen to speech naming one of the pictures. Does the infant look at the named picture?

iar target word 'baby'. Note that 15 month olds shifted their eyes to the picture of the baby *after* the target word was spoken. By 24 months, however, infants were several hundred milliseconds faster to respond, identifying the word 'baby' after hearing only the first syllable, *before* it was completely spoken. Other studies show that by 18 months infants are already becoming

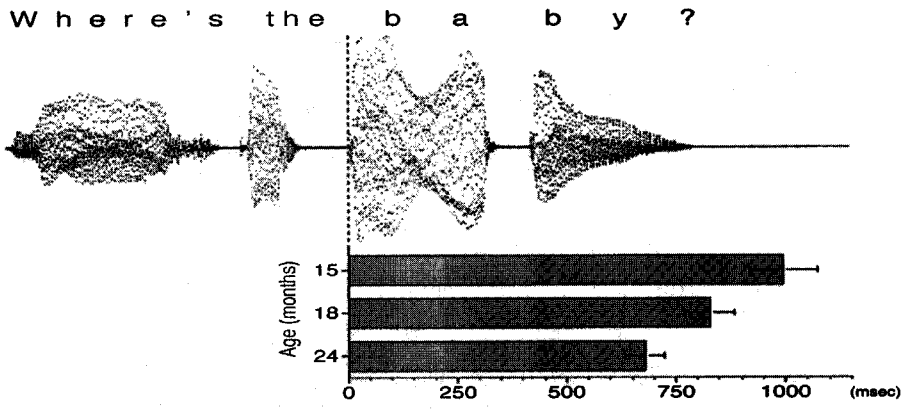

Figure 14.6 Mean reaction time to look at the named target picture when the infants were initially looking at the 'wrong' picture. Fifteen month olds take about a second (1000 milliseconds) and reaction time gets quicker with age. (From Fernald *et al.*, 1998.)

highly efficient in spoken language processing; like adults they can recognize familiar words based on incomplete acoustic information, a skill that is essential for the rapid and reliable processing of fluent speech (Fernald *et al.*, 2001).

Word learning and social understanding

Around the first birthday many infants begin to speak their first words, although some start a bit earlier and others don't talk for another 6 months (Fenson *et al.*, 1994). The first words spoken by most infants have a lot in common across different languages and cultures (Bloom, 1973; Casselli *et al.*, 1995). These include words used in social routines such as *bye-bye* and *peek-aboo*, names for family and pets such as *Mama* and *Fido*, lots of names for animals and common objects of interest to the child, such as *doggy, shoe*, and *train* (or *choo-choo*), along with a few action words such as *eat* and *push*. The infant's first utterances consist of single words, although some of these may be amalgams such as *gimmee*, which the child treats as a single word.

Many studies have investigated how children figure out the meaning of new words (P. Bloom, 2000). This might seem like a simple enough task: when the mother hands a new toy to the child and says 'This is a yo-yo', shouldn't it be obvious what *yo-yo* refers to? Actually it's not so obvious, because the new word could refer to many different features of the unfamiliar object. *Yo-yo* might mean *red*, or *round*, or *plastic*, or *string*, or *this-red-round-plastic-thing-with-a-string*. Questions about infants' cognitive strategies in guessing what new words refer to have been a research topic of considerable interest (e.g., Markman, 1989). But just as interesting are questions about how infants use their emerging knowledge of other minds to help them learn new words.

Emerging knowledge that others have minds

For example, we described earlier how children learn to follow the gaze of another person, and to understand that gaze can be indicative of that person's focus of attention. What role might gaze-following play in word learning? Considering the cluttered world most infants live in, adult gaze direction could be a valuable source of information to figure out what is being talked about. If an infant is playing with a fascinating toy car and hears his mother say 'Look at this yo-yo,' it would clearly be an error for the infant to assume that *yo-yo* referred to the car. Only by looking up to check and follow his mother's gaze could he make the correct association between the new word and the object it referred to. In fact, this is exactly what infants did in an experiment by Baldwin (1991). As the infant played with one toy, the adult said, 'It's a modi,' while looking at a second toy. At that point, most infants looked up to check the speaker's focus of attention. Later the children were shown both toys, the one they had been playing with and the one the speaker was looking at. When asked to choose the 'modi', 18 month olds chose the toy the speaker had been looking at, as if they understood that this was the one the speaker had *intended* to refer to.

As we have been stressing throughout, the ability to read the intentions of others has important implications for human communication. Someone who is unable to interpret social signals indicative of intent might therefore be expected to have difficulty learning to communicate. In fact, autistic children do seem to have trouble making use of social signals such as eye gaze, and, in an experimental study such as the Baldwin (1991) one described above, they also tend to make just the kind of mapping errors that non-autistic 18 month olds avoid (Baron-Cohen *et al.*, 1997).

Communication is very much a social skill, requiring the ability to interpret and understand the feelings, intentions, and mental states of others. Although infants can demonstrate an early understanding of some aspects of intentionality, it is worth emphasizing again that understanding intentionality is not an all-or-none ability. Carpenter *et al.* (1998) argue that prior to about 12 months, infants begin to recognize only *that* other people have intentions, although they are still unable to determine what those intentions are or how they come about. As a result, infants younger than 1 year can engage in activities where they share attention with another individual, but only as long as they do not have to follow another's attention or provide purposeful cues to direct another's attention. Only later, at about 12–13 months, do infants begin demonstrating an understanding of *what* other people's intentions are, which allows them both to follow another's attention and to consciously manipulate it. This new understanding enables word learning to begin in earnest, as infants can now treat a word uttered by another person as specifically intended to refer to a particular referent. Thus, in the Baldwin (1991) study described earlier, an 18 month old can recognize that a speaker intended to refer to one

Table 14.1 Average age at which infants first begin to demonstrate fundamental skills related to the emergence of communicative competence

Sharing of attentional focus with another person	9.0 months
Use of gestures to communicate	10.3 months
Following the attention of another person	11.5 months
Learning through imitation	11.9 months
Referential use of language	>15 months

(from Carpenter, Nagell, and Tomasello, 1998)

object even though the infant's own attention was focused on a different object.

Another impressive example of young children's ability to use a speaker's intentions to learn new words is provided by Akhtar and Tomasello (1996). In their study, 2 year olds participated in a finding game where they opened four containers and discovered an interesting object in each one. Importantly, these were objects that the children would not already have names for (e.g., a bike horn, a beanbag frisbee). After repeating this activity a few times, the children observed a speaker say 'Let's find the toma!' while unsuccessfully trying to remove one of the objects from a container that now appeared to be locked. Later, when the 2 year olds were shown the four objects and asked to select the 'toma,' they tended to select the object that had been in the locked container. Even though they had never seen that object labelled directly with the word 'toma,' they inferred from the speaker's unsuccessful attempts to open the locked container that he or she intended to refer to that object.

Conclusions

Human infants start out life capable only of involuntary signals closely tied to their physiological condition at the moment, not unlike other animals with a limited repertoire of inflexible social behaviours. From an evolutionary point of view, these primitive signalling behaviours serve a vital communicative function, because more mature and experienced humans are able to interpret them reliably as indices of the infant's needs and internal condition. But unlike other animals, the human infant's communicative skills expand and change continually from day to day, with experience and the development of new cognitive capacities that enable the child to extract socially relevant information from the voices, faces, and gestures of other people. By the end of the first year, the infant has moved well beyond the more limited capabilities of other primates, able to communicate intentions through symbolic gestures and words and to read the intentions of others through their social signals. During this

period the young infant demonstrates increasing knowledge of other minds through continual transitions in many dimensions—the transition from looking into the mother's eyes with delight to following the direction of her gaze for information, the transition from smiling in response to the mother's voice to calling her attention with a smile to something unexpected, the transition from babbling single syllables to using words symbolically to influence the behaviour of other people. Although a sophisticated understanding of others' thoughts, feelings, and intentions will still take several years to mature, the mind-reading ability that distinguishes humans from all other animals develops gradually over the first year, manifested first in non-verbal communication and then more clearly in the child's emerging ability to speak and understand language.

Introduction to

Infant development

EDITED BY

Alan Slater
University of Exeter

Michael Lewis
University of Medicine and Dentistry of New Jersey

2002

OXFORD
UNIVERSITY PRESS