14

Assisted imitation: affordances, effectivities, and the mirror system in early language development

Patricia Zukow-Goldring

14.1 Introduction

Rizzolatti and Arbib (1998) argue in their exposition of the Mirror System Hypothesis that brain mechanisms underlying human language abilities evolved from our non-human primate ancestors' ability to link self-generated actions and similar actions of others (see Arbib, Chapter 1, this volume). On this view, communicative gestures emerged eventually from a shared understanding that actions one makes oneself are indeed like those made by conspecifics. Thus, what the self knows can be enriched by an understanding of the actions and aims of others, and vice versa. From this perspective, the origins of language reside in behaviors not originally related to communication. That is, this common understanding of action sequences may provide a "missing link" to language.

In answering the question "What are the sources from outside the self that inform what the child knows?", the basic idea is that negotiating a shared understanding of action grounds what individuals know in common, including foregrounding the body's part in detecting that the actions of the self are "like the other." Given this footing, what then might the evolutionary path to language and the ontogeny of language in the child have in common? This perspective roots the source of the emergence of language in both as arising from perceiving and acting, leading to gesture, and eventually to speech.

I report here on an ongoing research program designed to investigate how perceiving and acting inform achieving a consensus or common understanding of ongoing events hypothesized to underlie communicating with language. This effort entails an analysis of the influences of the environment and, in particular, of the ways in which caregivers attune infants to that environment ("what the head is inside of": Mace, 1977). Building on what infants might "know" from birth, my work delineates the interplay of perceptual processes with action that might allow them to come to know "what everyone else already knows" (Schutz, 1962), including word meaning. Further, these caregiver practices may illuminate how automata might detect and learn new actions by observing and interacting with other intelligent agents.

Action to Language via the Mirror Neuron System, ed. Michael A. Arbib. Published by Cambridge University Press.

14.2 Affordances, effectivities, and assisted imitation: caregivers educate attention

The ability to imitate has profound implications for learning and communication as well as playing a crucial role in attempts to build upon the Mirror System Hypothesis (Iacoboni *et al*., 1999; Arbib, 2002). The focus here is to understand how imitation, especially assisted imitation, contributes to communicative development.

In the present study, I focus on how caregivers may assist infants as they gradually learn to adeptly engage in new activities that they have observed others doing. This interactive process may provide as well the basis for a shared understanding of events. The key notions for this study are "affordances" and "effectivities." J. J. Gibson (1979) proposed the notion of *affordances*, referring to the ability of creatures to perceive opportunities for action in their environment. The classic example is that as we walk across a room we see more, and that the more that we see tells us which surfaces will support our walking, what objects block our way, and so on. Some characterize the relation as emergent properties of the animal–environment system in which affordances are specified relative to an agent (Stoffregen, 2003), others add the notion of *effectivities* – the repertoire of what the body can do – as a dual complement of affordances (Shaw and Turvey, 1981; Turvey *et al*., 1981; Turvey, 1992). I suggest that effectivities expand as an individual gains skill participating in new activities, thus concomitantly differentiating further what the environment affords for action.

The question, however, is how does an infant become more adept? What to do with objects, beyond the most rudimentary self-directed actions such as sucking and grasping, presents a challenge. Can infants, novice members of their culture, detect affordances and consummate an activity without assistance, if the action is not present at birth? (See Oztop, Arbib, and Bradley, this volume, for the development of grasping that is present at birth.) I have argued that objects cannot "tell us" what they afford (Zukow-Goldring, 1997). Nor can caregivers of young infants *tell them* in so many words, since verbal instructions directed to infants before they know "what and that" words mean, prove most ineffective (Zukow-Goldring, 1996, 2001). Novices do learn affordances as they engage in daily life in a particular time, place, and culture (Zukow, 1990; Costall, 1995). However, even careful non-verbal guidance – pointing out aspects of elements, configuring them in time/space, and/or modeling actions – may not be sufficient, because the body's work is left implicit. Shaw (2001) made that work explicit. He has argued that effectivities transform potential experiences into actual ones; that is, *affordances dispose*, while *effectivities deliver* (actualize). As Kadar and Shaw (2000, p.161) suggest, the fit between affordance and effectivity "is not between equals, but between promise and performance." This position suggests that affordances will go unrealized and, perhaps, undetected unless the creature has the requisite bodily ability within its repertoire. Elaborating further, Witherington (2005) argues that the range of what affordances offer an individual is not available to a creature prior to interaction with a particular object; rather the novice discovers them in and through agent–environment coupling. Therefore,

in contrast to other views of imitation, I suggest that unless the novice perceives or knows the "body's part," affordances will not be picked up.

I propose that the child at first lacks the ability to detect by observation alone how the body relates to the physical layout and to the furniture of the world, except for the most rudimentary actions. From an ecological realist view, we might say that directing attention facilitates detecting a field of practice in which perceptual learning can take place. E. J. Gibson (1969) investigated the development of affordances present at birth in terms of perceptual differentiation. Studies of such affordances are important for understanding pre-attunements underlying basic modes of functioning as well as the subsequent role of experience. However, E. J. Gibson and Rader (1979) noted that most human affordances are learned, stressing observation of others over instruction by them (E. J. Gibson, 2002). Elaborating, E. J. Gibson and Pick (2000) discussed how the social context "acts back," responding to the infants' spontaneous exploratory activity, but did not discuss whether or how caregiver and infant interact in the learning of affordances. Filling in the gap, I stress that the caregiver *educates attention*. She guides the infant to discover for herself the means to realize the aim of an action (its goal) through the inextricable link between detecting the perceptual structure specifying affordances and the gradual differentiation of her own effectivities.

What do infants have to learn about the world? Infants must learn the most basic things, e.g., taking a bath, eating with utensils, walking. During mundane activities, infants must detect and participate in assembling the structure and organization of everyday events before they can communicate with others about these events (Zukow, 1989). In part, the child can discover affordances and effectivities as she explores the world on her own, developing a useful repertoire of skills by trial and error (Oztop *et al*., 2004). However, I stress the role of the caregiver in directing the child's attention in ways that greatly expand this basic repertoire of affordances and effectivities, potentially reducing the search space and thus speeding learning. In contrast, many studies and theories assume that children know and/or learn autonomously how their bodies move in space and in relation to animate and inanimate things (Piaget, 1962; Thelen and Smith, 1994) and thus do not explore what experiences might underlie eventual adept performance. Most research investigating the development and implications of imitation ("learning to do something from seeing it done": Thorndike, 1898), focuses on what the child knows, rather than how the child comes to know.

Accounting for these achievements usually takes the form of proposing some combination of maturing modules, socio-pragmatic knowledge, or cognitive precursors hypothesized to be necessary for the activity (Uzgiris, 1991; Tomasello *et al*., 1993; Meltzoff and Moore, 1995). This literature documents the age at which the average child can observe someone else's action and repeat it accurately either promptly or after a delay. The actions investigated, such as tongue protrusion or grasping, transferring, and stacking objects, are part of the child's repertoire. What is novel may be the object acted on and/or the sequence of acts, not the action itself. This literature may underestimate sources of the infants' accomplishments located in the caregiving environment. Greenfield (1972, this

volume) observed that children imitate those actions that are entering their repertoire. Why might these particular actions be ripe for imitation and not others? Are the children's imitations usually autonomous accomplishments or do they have a robust history of assistance from others? Piaget (1962, p.18) asserted that only the infant's independent achievements contribute to cognitive development. In the same vein, he referred to the "pedagogical mania" of those who tend children as interfering at best. Yet, as many researchers have documented, learning a new skill by observation alone can be a very slow, trial-and-error process whether in human (Vygotsky, 1978), robot (McGovern and Barto, 2001), or non-human primate (K. R. Gibson, 1993). Fortuitously, caregivers do invite infants to imitate, and I suggest that this is for the better. *Assisted imitation*, informed by an integrative view of action and perception, delineates how educating attention may contribute to achieving a consensus or common understanding of events hypothesized to be a prerequisite for communicating with language.

I focus on learning a new activity (not in the infant's repertoire) which may not be easily attained by observation alone. However, Dautenhahn and Nehaniv (2002a) asserted that the possibility of confirming whether or not an activity/behavior is in an intelligent agents' repertoire is very problematic, because no one has a complete history of someone's actions from birth. They claim that "new" is always a matter of degree. Nevertheless, an infant's behavior often can and does tell those who know them best how to interpret what the infant does or does not know. The infant's behavior and the caregiver's response to that behavior display how the caregiver evaluates the infant's current level of skill. In particular, if an infant initially misunderstands a caregiver's messages inviting her to do something, the caregiver has evidence that the infant does not know how to respond, i.e., that the behavior is not in the infant's repertoire (Zukow-Goldring, 1996, 2001).

The key idea here is *assisted imitation*: I argue that caregiver practices guide infants to perceive possibilities for action (Zukow-Goldring, 1997).[1] Out of the unceasing perceptual flow, which is quite unlike the highly edited cuts of most movies, *caregivers* continuously educate attention to aspects of ongoing events. Further, our normal experience is highly multisensory, not restricted to the limited perceptual input of, say, a videoclip. Indeed, Stoffregen and Bardy (2001) have argued that multisensory perception is not merely the primary type of perception; it's the only type of perception. Caregivers and children detect "the something that something is happening to" as well as "the something that is happening" through vision, smell, hearing, taste, movement, and touch (Michaels and Carello, 1981; Zukow-Goldring, 1997). Especially relevant to this idea is the young infant's known ability to detect amodal regularities or invariants in the continuous stream of perceptual information (Spelke, 1979; Bahrick and Pickens, 1994). Caregivers guide

[1] In the same vein, Ingold (2000, p.190) has documented how the perceptually more adept make what they know available to the less skilled, guiding the attention of others along the same paths.

infants to notice key elements of what persists or remains invariant over time and what changes.

In what follows, I illustrate the findings from a number of studies of infant development with some qualitative examples (Zukow-Goldring, 1996, 1997, 2001).

14.3 The naturalistic investigations

The infant's initially unsuccessful attempts at imitation with a toy or food item often display some familiarity with the cultural use of objects, as she attempts to engage in the relevant sequence of activities. Grasping an object is usually the fulcrum around which novel action grows. Even though such objects and their uses are not entirely novel, what is required to imitate apparently is. That is, the ability to notice the relevant affordances and coordinate them with particular effectivities that are necessary to accomplish these tasks is not available to the infant without assistance. Infants' fragmentary, flawed attempts to imitate actions observed in the past elicit very careful and elaborate tutoring on the part of the caregivers to direct attention to relevant affordances and effectivities. Going further, we need to understand how picking up the perceptual information that the caregiver has highlighted allows the infant to get a grip on what to do. Getting a feel for what to do can provide the basis for detecting the affordances that will guide infants in their attempts to imitate that action.

In these studies the following hypotheses were tested:

(1) Providing a child with more perceptual structure specifying either effectivities or affordances will assist caregiver and child to achieve the consensus or common understanding needed for communication, including, where appropriate, explicit guidance of the child's movements.
(2) Additional or more specific verbal information will not enhance understanding when no basis for that understanding has been embodied.

The examples from the studies reported illuminate how a human child learns about the world. Of course, I do not deny the utility of verbal instruction for older children or their own autonomous learning. Rather, the purpose is to illustrate how the fundamental link between perception and action provides the information upon which communication can build. It is a separate study to understand the later "bootstrapping" that occurs when words can take a far greater role in advancing what the child knows.

14.3.1 Sample and data collection

Five Euro-American middle-class families and six Latino working-class families with an infant of 6 months were followed monthly through the one-word period. Monthly 20-minute videos of naturalistic interaction at home, field notes, diaries, and check-lists of lexical development were collected, as well as interviews following each videotaping session to ascertain the caregiver's interpretation of ongoing events and of the infant's utterances.

14.3.2 Procedure: attention-directing interactions

Situations were selected in which caregivers directed infants to notice the content of messages, such as specific elements, relations, or events over the myriad other possibilities available. This collection of *attention-directing interactions* included all instances of perceptual imperatives expressed by caregivers, such as *look!*/*¡mira!, listen!*/*¡oye!*, and so on, as well as the accompanying gestures, and the gestures alone as well as the infants' subsequent actions. Zukow-Goldring noticed this set of perceptual imperatives when doing fieldwork in Mexico (1981–2). They occurred massively. Caregivers constantly said, *¡mira!*/*look!*. Surely there is no *mira* neuron; however, the use of perceptual imperatives may help draw the child's attention to the gestures that co-occur with them. These gestures may, in turn, provide specific support for imitation by directing attention to the perceptual information (in touch, smell, taste, vision, movement, and hearing) that may lay the groundwork for knowing that the self is like others. Thus, this ability opens up the possibility of learning the actions that others display.

A key issue was to determine the relative importance for the young child of verbal messages versus gestures that educate attention and action. The distinction here is between words that provide explicit instructions, such as *Peel the orange!*, and the pairing of perceptual imperatives like *¡mira!* and gestures that direct the child to attend to opportunities for action during ongoing events. The conclusion was that explicit verbal instructions are ineffective in the early stages. (For further details, see Zukow-Goldring, 1996, 1997, 2001.)

Targets of attention

Caregiver messages combine gestures with *targets of attention* throughout the prelinguistic and one-word periods. In messages, caregivers express what persists and changes as events coalesce and disperse. The targets vary in semantic complexity, which we categorized in four levels: non-dynamic animate beings or inanimate objects; simple action or change of state of the former; more complex relations within an event (Greenfield and Smith, 1976); and relations between events. These targets of attention bring to life structural and transformational invariants in the environment across space/time.

Attention-directing gestures

Five gestures that direct attention often accompany caregivers' verbal messages (Fig. 14.1). These gestures direct attention to perceptual information/semantic content through action. The gestures encompass varying degrees of other- to self-regulation of attention to the effectivities of the body and the affordances of the environment. When a caregiver *embodies* his infant, he puts her through the motions of some activity, so the two move as one (e.g., the caregiver takes the child's hand, using his own hand on top of the child's to press down a lever as he says, *¡por abajo!*/*down!*). The infant may already "know" the movement, but this activity calls for a new way to match or fit the effectivity of the hand/arm movement to the affordance of an object or the action

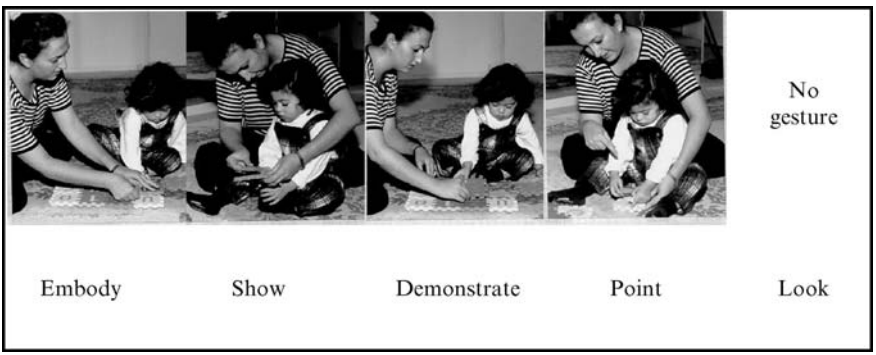

Figure 14.1 Attention-directing gestures. The child was about 2 years old, an age at which some children approach the end of the one-word period.

may be completely novel. Intriguingly, an inanimate object (like the spring of the vibrating toy described below) can also *embody* in this sense by putting the child "through the motions" in some way. During *shows*, a caregiver regulates the infant's line of sight. Sometimes the path of a gesture generated by a translational motion of an object through interactional space successively occludes other information in a scene or can magnify detail by bringing the object closer. For instance, the caregiver looms a toy dog toward the infant, saying *gwow-wow*/*wow-wow*. Other times, the caregiver performs some action using a familiar bodily effectivity to introduce a new possibility for action with an unfamiliar object or affordance. For instance, she pushes a button on a new toy, saying *¡empújalo!*/*push it!* The caregiver does not request that or leave interactional space for the infant to join in, but moves on to some subsequent activity. In contrast to *shows*, during *demonstrations* an infant is invited to act/imitate through action and/or speech *(¿Y tu?*/*You do it!)*. The infant who watches closely must detect or pick up in the perceptual flow a familiar coupling of effectivity and affordance to be duplicated. For example, the caregiver may synchronize rhythmically retracting fingers of an upright palm with saying *adiocito*/*bye-bye* when catching gaze and smiling. Alternately, an infant may be asked to pretend to avoid the sharp spines of a prickly pear while the caregiver mimes approaching and pulling away from the fruit's surface, saying, *¡espinoso!*/*prickly!* Regarding *points*, the infant must detect where a gesture's trajectory through space converges with some target of attention (the caregiver pointing to and saying *p'acá*/*over here*). For *looks*, no gestures accompany the caregiver's speech. Instead, only the caregiver's words and gaze direct the infant to correlate attention with that of the caregiver.

Assessing consensus or shared understanding

We determined whether or not caregivers treated their infants' responses to each message in a sequence as adequate or not. Action or speech indicating approval or embellishing the ongoing activity followed acceptable attempts (understanding), whereas inadequate

responses (continued misunderstanding) were followed by repeated and revised messages, terminating the current activity, or noticing the child's lack of interest. To assess *perceptual structure*, each message in a sequence after the initial one was scored as providing or failing to offer additional perceptual information. To evaluate *linguistic specificity*, each sequence, message by message, was assessed for increases or decreases in linguistic specificity. Linguistic messages in ensuing turns can contain more or less explicit expression of nouns and verbs and previously ellipted/omitted lexical items. For instance, when adding linguistic specificity, the caregiver might say *Peel it!* and then, *You peel the orange!* Caregivers also can express less specificity in subsequent messages, e.g., *Peel it!* followed by *Do it!*

14.3.3 Qualitative examples: assisted imitation

Caregivers regularly take care to arrange the physical layout, so that the configuration in space of caregiver, infant, and object(s) makes them suitably aligned so that action is within reach. In addition, optimal proximity makes perceptually prominent what the object or some aspect of it affords for action. Sometimes the caregiver *demonstrates* the action(s) and then gives the infant a chance to act. Quite often the infant's attempt is inadequate. Frequently the caregiver embodies the infant, so that the child can perceive the relation of his or her body in terms of posture, motor actions, and rhythm to the caregiver as they move as one to accomplish the action or action sequence. Of the three examples, the middle one focuses less on tutoring effectivities and affordances and more on learning a sequence of actions to consummate an activity (vibrating toy).

Pop beads (13 months): caregiver tutoring of effectivities and affordances when concatenating beads

Pop beads, easily graspable by infants and toddlers, have affordances that allow concatenation. Play with this toy consists of (i) orienting toward each other the parts of each bead that afford concatenation (the dual complements of protrusion and opening), (ii) moving the appropriately oriented beads toward each other on a converging path, and (iii) applying enough force when the parts meet to embed the protrusion of one in the opening of the other.

The infant, Angela, begins by pressing a block lacking the appropriate affordances and a pop bead together. She displays an understanding that completing the task requires two small graspable objects and the application of some force to bring them together (Fig. 14.2, PB1). Her behavior provides no evidence that she knows that a set of objects with specific parts must come together, nor that they must sustain an orientation as they meet along a converging path. Her mother, Cecilia, provides perceptual information to Angela, gradually foregrounding the affordances of the objects and the effectivities of the body required to put the beads together. At first, Cecilia provides a bit of both, *point-touching* the opening on one bead (but not the protrusion in the other) as she directs

Figure 14.2 Pop beads.

attention to an affordance (PB2) and then reorients another bead (PB3), enacting a movement that aligns the beads on the same converging path. Lacking at this point, however, is information displaying the path itself or the force required to push the converging objects together. After an unsuccessful attempt by Angela, Cecilia *shows* her what the body must do to move the beads along a path with the required orientation as she pushes the protrusion into the opening with appropriate force (PB4). The infant

remains unable to put the beads together (PB5) The mother then more elaborately provides perceptual information as she slowly *point-touches* both the affordances of protrusion (PB6) and opening (PB7), followed by *demonstrating* the effectivities required for connecting and then disconnecting of the beads (PB8). As the infant watches more intently, Cecilia eventually invites Angela to imitate: *¿A ver, tu?/Let's see, you (do it)?* She assists her daughter's imitation by *partially demonstrating* what to do by orienting a bead opening toward Angela and holding it in a fixed position as she makes the protrusion easy to see on the second (affordances for action) (PB9). Angela moves the bead along the appropriate path (PB10), but she misorients her bead's protrusion from the opening of the other as the beads touch one another (PB11). Cecilia realigns the opening of her bead, making prominent just where Angela should push in her bead. As the infant pushes in the protruding end of her bead, the mother pushes from the opposite direction with enough force to link the beads (PB12).

In this case, the caregiver's gestures gradually provided increments in perceptual information that guided the infant to concatenate two objects. Eventually, the caregiver simplified the task by holding an appropriately oriented bead as a fixed target. This assistance allowed the child to bring her slightly misoriented bead (i) along a path toward her mother's (ii). Angela pushed her bead against the other (iii), as her mother subtly reoriented her bead (i) and provided a complimentary push (iii).

Notwithstanding Angela's noteworthy improvement on this occasion, bringing together two hands, each grasping a properly oriented object, was not within her "reach." Nor could she by herself apply enough force to connect her beads. It is possible that embodying Angela, putting her through the motions, might have drawn maximal attention to the inextricably linked affordances of the beads and effectivities of the body required to consummate this activity. The point is that simply observing someone engage in a complex activity does not prepare the young child to imitate it. Much tutoring is required to build a repertoire on which "true" imitation (if such exists!) can take place.

Vibrating toy (14.5 months): caregiver and "toy" tutoring of a sequence of actions

This infant, Elsa, and her mother, Kathy, engage in a familiar routine with a reindeer toy that has a hidden affordance, a spring inside the toy to which a string is attached. In this routine, when the caregiver pulls on a ring that protrudes from the back of the toy, the string unwinds. Releasing the ring/string at the apex of its extension retracts the string so quickly that the toy vibrates strongly accompanied by a loud pulsing noise. Elsa expresses delight when she feels the vibrating toy placed on her stomach. Elsa, however, cannot make the toy vibrate by herself. This "game" entails a sequence of actions: (i) someone grasps the string by the ring, (ii) the string unwinds as it is pulled, and (iii) retracts within the toy as the tension on the string lessens. Finally, (iv) someone places the vibrating toy on the infant's stomach.

Family members had played this "game" with Elsa quite frequently during the prior 11 months. However, she had never attempted to imitate the others. In this example, more emphasis is placed on the sequence of actions than on the briefly noted affordances

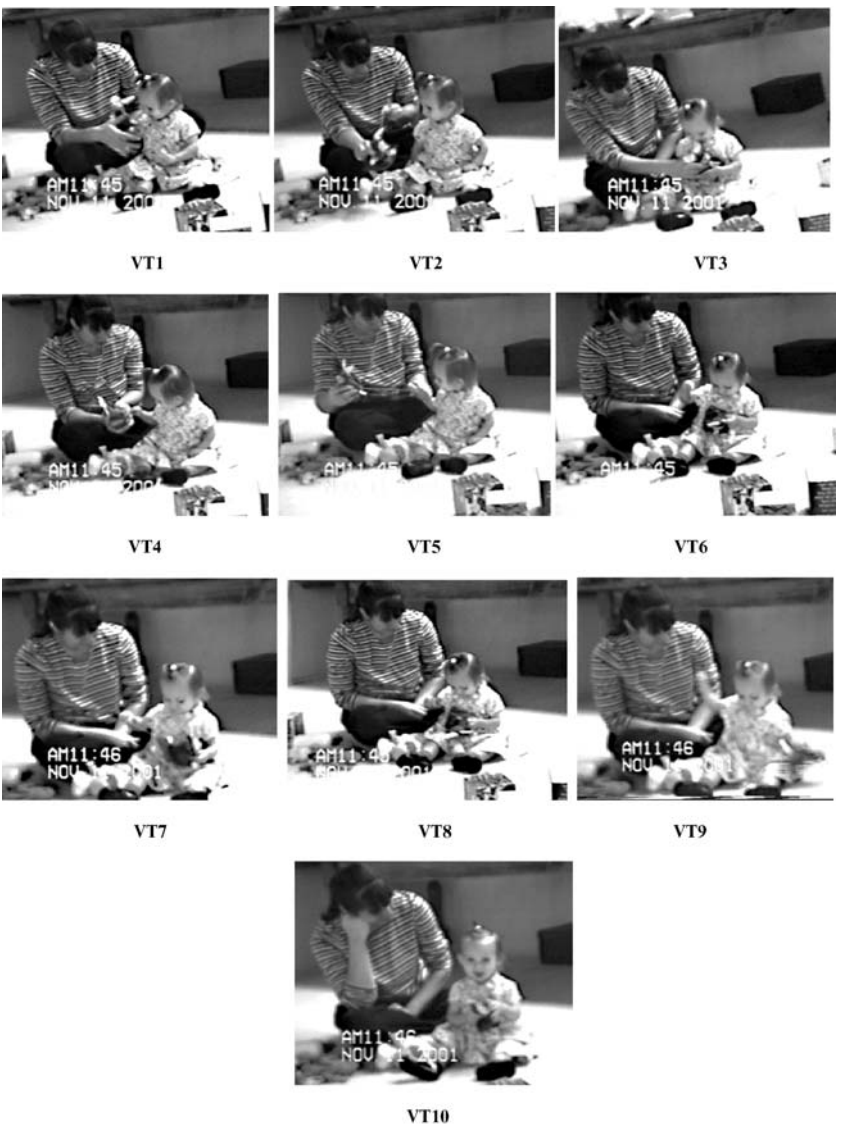

Figure 14.3 Vibrating toy.

and effectivities that mother and toy make perceptually available. For conciseness' sake, this sequence has been abbreviated by omitting repetitions and variations leading to the child's final adept enactment of this activity.

Elsa, sitting in front of her mother, turns to give her mother, Kathy, the toy that she wants her to animate (Fig. 14.3, VT1). Kathy pulls the string out by the ring (VT2), releases the string, and places the vibrating toy on Elsa's stomach (VT3). When Elsa wants her mother to continue, Kathy says *You do it!*, as she *partially demonstrates* by

orienting the back of the toy toward Elsa, making the ring for pulling (affordance) prominent and within the infant's reach. The infant grasps the ring (VT4). The mother *embodies* by putting her hand/fingers on top of Elsa's ring-grasping hand to hold her hand steady as she pulls the toy away from them, presumably so the infant can feel the tension (affordance) of the string as it unwinds (VT5). The spring attached to the string *embodies* Elsa by pulling her hand, as it holds the string, back toward the toy (VT6). That is, the two move as one as the contracting spring pulls both Elsa and the string along the same path. Subsequently, the infant pulls the string out herself (VT7) and holds on as the spring *embodies* by retracting the hand holding the string back toward the toy (VT8). At this point, the toy vibrates weakly, if at all. A few seconds later, Elsa pulls the string so quickly and fully out, that the tension on the string *embodies* her by snapping the ring from her fingers. Vigorously pulling the string (effectivity) allows her quite serendipitously to experience how to take advantage of the vibratory properties of the toy (affordance). Note that her arm recoils from the force moving quickly away from its former position, while the hand holding the strongly vibrating toy moves far in the opposite direction (VT9). Within seconds, Elsa first pulls and lets go of the string and then places the base of the toy on her stomach to best perceive the vibrations as she expresses evident joy (VT10).

Note the free building-up of a sequence, as the child understands each new element. Both caregiver and toy educate Elsa's attention, cultivating perceptual differentiation of new affordances and the refining of her actions (effectivities) to fill in the gap between grasping the ring and feeling the vibration of the toy on her stomach. Elsa experiences bodily the tension of the string unwinding as her mother pulls the toy away from her and as the spring hidden in the toy that controls the string retracts as the string appears and disappears. Pulling the string out is within Elsa's grasp. However, the accidental snapping back (letting go) of the string at the apex of its path and at its highest tension made evident the relation between the effectivity of releasing the string and the ensuing affordance of the toy's vibrations (see Zukow-Goldring and Arbib (in press) for a fuller analysis of this sequence). By the second attempt, Elsa had placed the toy to bring the most enjoyment. Within the next several minutes, she changed her grasp from a finger crooked through the ring to a pincer grip, could pull the string with both right and left hands, and attempted to give her doll the same experience. Although she knew "that" the toy held potential for vibrating, she did not know "how" to make it happen until she received very careful tutoring. This educating of attention and action contrasts sharply with the effort it takes to tutor the attention and action of monkeys as well as autistic children (Nadel *et al*., 1999; L. Fogassi, personal communication).

Orange peeling (16 months): caregiver tutoring "when actions speak louder than words"

Peeling an orange with the hand entails penetrating the peel (both zest and pith), grasping the pulled away edge in a pincer grip, pulling the peel away from the flesh, and separating that portion of the peel from the fruit. Tearing off a piece of peel may involve yanking

Figure 14.4 Orange peeling. (a) Hand, head, and finger points and increasingly specific verbal messages do not convey just how child and orange engage in "you peel it." Cecilia embodies her child to provide the missing perceptual information. (b) Cecilia partially demonstrates by lifting the peel almost entirely free of the flesh. She invites Angela to join in, saying *A ver, estírale*/*Let's see, pull it*. Now, Angela knows how. (c) Cecilia makes the task more difficult, peeling only part way through the pith. Angela cannot tear through the zest. Cecilia mimes "strong and hard" *as well as* change in direction. Angela imitates, but still needs help.

it away or, if the zest is tough, rotating the wrist at a 90° angle from the fruit. Thus, what to an adult seems like a single action must for the child be learned as a sequence of grasps and directed actions.

In the first vignette, the infant, who is quite fond of oranges, has seen them peeled and wants to do so herself, but cannot. She scratches quite ineffectually at the surface of the peel, but does not know how to remove it. Angela needs her mother's help to learn the actions that go into peeling an orange. However, increasingly explicit verbal messages (*Let's see. Peel it! You peel! Take it off like this, look!*) coordinated with a variety of points, ranging from a subtle head point to proximal points of hand and index finger, do not communicate just how to peel an orange. In contrast, *embodying* the child provides the missing perceptual information (Fig. 14.4a) regarding how effectivities of her body and the affordances of the orange continuously inform one another. The mother's hands shadow and guide those of the child as together they each support the orange with one hand and pull off the peel with the other as they move together. *Embodying* the infant provides her with perceptual information in vision, touch, and movement that she is like her mother (the other). (For more detail, see Zukow-Goldring, 2001.)

In the second vignette (Fig. 14.4b), Cecilia *partially demonstrates* by pulling the peel nearly free of the orange. Angela easily removes that bit of peeling.

In vignette three (Fig. 14.4c), Cecilia ups the ante as she lifts the peel just a small distance away from the flesh. In the third frame above the mother mimes what she says *¡Duro–duro–duro! ¡Fuerte–fuerte–fuerte!/Hard–hard–hard! Forcefully–forcefully–forcefully!* However, she provides perceptual information for both manner of action (hard, forcefully) and trajectory (a change in direction of about 90°) in gesture, but only for manner in speech. Note that trajectory is the information that helps solve the problem, illustrating when and that gestures "speak" louder than words (Kendon, 2004). Although Angela changes the direction of her pulling about 45° (compare frames 2 and 5), Cecilia helps by *showing* how to pull away the peel at a more extreme angle.

14.3.4 The naturalistic investigations in perspective

I have argued that ambiguity is the rule, not the exception. Words cannot explain unless a person already knows what words mean. To understand what words mean a person must understand what is happening. Novices do not automatically understand the organization and structure of daily events. Experts show them. These qualitative examples emphasize the finding that caregivers work hard to cultivate their infants' understanding of ongoing events by providing perceptual information in the form of attention-directing interactions to disambiguate their initially misunderstood messages. When there is lack of consensus, providing more perceptual structure tends to resolve misunderstanding, whereas adding specificity to verbal messages does not reduce ambiguity. Further, as infants respond to caregiver guidance, the infants' misdoings that may ensue pinpoint possible misperceiving as well as lack of bodily skill that, in turn, may inform the caregiver's subsequent seeing and doing. Perceptual restructuring

of messages following communicative breakdowns led to achieving a common understanding in both working-class Spanish-speaking and middle-class English-speaking families in the United States.

From being a body to becoming a cultural being "like the other"

Some attention-directing interactions may have another equally important function: caregivers foreground the correspondence between the effectivities of the infant-actor's body and that of their own, providing an opportunity to learn the self is "like the other." In contrast, Meltzoff and Moore (1995, 1999) and many others propose that infants already know that others are "like me," giving the infant the basis for imitating.[2] Instead, I suggest the reverse. As the infant knows relatively little, another "like me" would be of little assistance. Knowing the self is "like the other," on the other hand, would provide access to more skillful behavior. When caregivers *embody* their infants, the infants have a chance to see and feel that their own movements are "like the other." That is, as caregiver and child move as one, the infant can detect amodal invariants (proprioceptively, kinesthetically, visually, tactilely, and so on) in the synchronous onset/offset, rhythm, and tempo of the action that may specify the correspondence between caregiver and infant movements. Embodying experiences may provide the basis for the infant's later ability to benefit from *shows* and *demonstrations* as the infant must already know that the self can move "like the other."

Byrne (2003) has drawn attention to the fact that scholars use the term *imitation* in two ways:

- *the transfer of skill problem* (how can the child acquire *novel, complex behavior* by observing?)
- *the correspondence problem* (how can the child match *observed actions* with *self-executed actions?*).

Thus, if a child knows that she is herself "like the other" (e.g., the caregiver), perhaps she can learn to do what the other does to achieve similar benefits or avoid risks. But can a novice (who does not know she is "like the other") spontaneously or after a delay imitate just any novel and/or complex "developmentally appropriate" behavior not yet in her repertoire without assistance? From a correspondence view of imitation (knowing the other is "like me"), the expectation would be that the novice can do so (Meltzoff and Moore, 1995, 1999). However, examples of *assisted imitation* provide evidence that the answer is "probably not." Instead, I suggest that investigating the *transfer of skill problem* can provide answers to the *correspondence problem* as well. That is, understanding the caregiver methods that promote the *transfer of skill* can illuminate how the infant comes to know the self is "like the other." Thus, as the child learns new bodily skills, she may literally get in touch with both sides of the *correspondence problem* perceiving the match between the other's actions and those of the self. (See Oztop and Arbib (2002) and Oztop *et al*. (this volume) for modeling of the development of grasp-related mirror neurons,

[2] For a cogent critique of prevailing cognitive approaches, see Moore's (2005) alternative cognitive proposal explaining how infants achieve an understanding of "self-other equivalence."

suggesting how the mirror system grounds imitation as a core component of communicative and linguistic development within an action/perception framework.) I stress again the role of the caregiver in directing attention to *effectivities* as well as *affordances* of the self and others – the two sides of the mirror system.

Challenges for the future

A challenge to these findings is the following question: do caregivers from all walks of life across the world's cultures engage in these methods on a daily basis? Evidence from studies in Indonesia, England, and France (Bril, 2004), the Solomon Islands (Watson-Gegeo and Gegeo, 1986, 1999), the Maya of Mexico (Maynard, 2002), rural and technological sites in Mexico (Zukow, 1989), and the Marquesas (Martini and Kirkpatrick, 1992) suggest that a broad range of both sibling and adult caregivers engage in very similar imitative sequences. (For a brief treatment of similar caregiving received by the deaf–blind Helen Keller, see Zukow-Goldring, 2001.)

An additional question entails whether or not "pedagogical mania" makes a difference: do infants who engage in assisted imitation most frequently develop communicative skills more rapidly than those who do not? Future research can reveal the answer to this question and enrich our understanding of these phenomena. However, sufficient may be a more likely answer than "most." In contrast, evidence from a clinical case study suggests that too little of these caregiving practices apparently is detrimental. In the case of a retarded caregiver who did not carefully coordinate speech and actions in terms of content and timing, her cognitively normal children displayed arrested language abilities (N. Rader, personal communication).

In sum, caregivers establish an understanding of what is happening. They gather and direct attention to perceptual structure that makes prominent the relations among animate beings, objects, and their actions. These dynamic relations specify the organization and structure of the most mundane daily activities. Caregivers introduce their infants to new effectivities or bodily capabilities and affordances for action and interaction on a daily basis. They assist them to link sequences of actions that comprise more and more complex activities. Caregivers also set aside language training when communication breaks down and, instead, focus on providing the perceptual information that will lead to a consensus. As caregivers educate attention, infants gradually learn to perceive, act, and know in culturally relevant ways.

14.4 Child development and the development of imitation in robots

Before turning to the discussion, the relation of this naturalistic research to developmental robotics is explored. Zukow-Goldring and Arbib (in press) have looked at the implications of these findings, structured around a number of questions raised by Breazeal and Scassellati (2002).

How do robots know what to imitate? While much work in robotics involves explicit programming of robot behavior, Zukow-Goldring and Arbib's concern with "robot

development and learning" focuses attention on studies in which researchers prime robots to monitor preselected, simplified aspects of the visual scene, such as certain colors, areas of two-dimensional space, human movement, and so on to ensure the identification of "what to do" (Billard, 2002). Goga and Billard (this volume) offer an explicit model of caregiver–child interaction, using artificial neural networks to control the behavior of two robots, but here the "child" simply attends to the actions of the "caregiver," without specific attentional guidance on the part of the caregiver. However, much work in developmental psychology has established that sharing the target of attention during interaction plays a crucial role in cognitive and language development (Bates, 1976; Bruner, 1983; Zukow, 1990; Moore and Dunham, 1995). Rather than focusing on what the child already knows how to do, some of these scholars investigate the conditions under which infants learn new behaviors during social interaction. Several developmental approaches explicitly examine interactive processes that can lead a novice to learn a new task, including Vygotsky's (1978) groundbreaking work on the zone of proximal development, followed by investigations of scaffolding (Wood *et al*., 1976), other-regulation (Wertsch *et al*., 1980), and fine-tuning (Snow *et al*., 1987). This research investigated what children need to know about objects of action and how caregivers guide them with verbal and non-verbal messages. However, the body's work is left implicit. In contrast, during assisted imitation, unless the individual perceives or knows the "body's part," affordances will not be picked up.

How do robots map the perception (of other's actions) onto their own action repertoire to replicate it? Dautenhahn and Nehaniv (2002; Nehaniv and Dautenhahn, 2002) suggest that solving the correspondence problem (matching *observed actions* to the *self's actions*) answers this question. For instance, Atkeson and Schaal (1997) have endowed automata with a priori knowledge of the goal of some action, the means to match movements observed to movements in the robot's repertoire, and provided reinforcement schedules that resulted in rapid matching of behaviors (see Schaal, this volume). However, children are not so endowed from birth, but must learn the aim of an action as well as the means, e.g., movements that result in achieving the target activity.

Bearing on these issues, Breazeal and Scassellati (2002) contrast descriptive studies of what biological agents can do and when with generative investigations of artificial systems. In such studies, actions are built up piece by piece, a very complex problem for those who design automata. The authors, however, do not make explicit in their critique that many studies of biological agents take the body's work for granted. In the main, researchers assume that the creature has in its repertoire whatever movements are necessary to imitate the activity observed. However, cultural knowledge of object use is not present at birth; such knowing must be learned. As studies of assisted imitation demonstrate, infant novices must learn to use their bodies in new ways as behaviors emerge during interaction with the environment.

Dautenhahn and Nehaniv (2002a) and Schaal (1999) have called for addressing imitation from new perspectives, emphasizing the import of perception–action coupling to the creation of autonomous humanoid robots. Exchanging skills entails at least two embodied

agents in a particular environment who display information to each other via verbal and/or non-verbal communications. In one use of forward models,[3] the agent can interpret the visual scene by anticipating the consequences of her/his actions and so guide ensuing action toward that end. Oztop *et al*. (this volume) discuss the relevance of forward and inverse models to the operation of the mirror system, and Skipper, Nusbaum, and Small (this volume) carry this theme into the study of brain mechanisms for audiovisual speech production and recognition. From a computational perspective, the success of training results from the child's ability to develop a forward model (Jordan and Rumelhart, 1992) that can link the (unobservable) internal processes to their observable consequences. The solution must relate to developing a means to discover the goals that characterize the success of diverse actions. From this computational perspective, one might say that the caregiver draws the "infant's" attention to the affordances in the environment to provide goals for the infant's/robot's developing effectivities. From an ecological realist view, we would say that directing attention provides practice, so that perceptual learning or differentiation can take place. Thus, whether in infant or robot, the actual specification of the processes guiding the learner are beyond the reach of even the most explicit instruction.

Zukow-Goldring and Arbib (in press) stressed the importance of *assisted imitation*, and the utility of the caregiver *embodying* the infant. With a robot, the challenge is greater since one can no longer guarantee so direct a map of action–perception coupling between "caregiver" and robot as there is between the adult human and child. Thus the robot situation poses extra challenges in finding the novel effectivities that match the affordances to which the caregiver draws attention. However, with robot or child, the skill of the caregiver resides in assisting the novice to detect the inseparable coupling of its/her own effectivities with the affordances of the environment.

Clearly the robots to be studied in future developmental robotics are very different from the highly programmed robots of today's assembly lines. However, note an interesting divergence between developmental robots used as models of child development and those being prepared for application in the workplace. In the former, the exploration of affordances and effectivities provides the basic substructure for language. In the latter, basic symbol structures can readily be programmed on the underlying hardware so that directing a robot's attention to a specific part of its body can be done symbolically in a way closed to the infant.

What would be gained if an instructor notices the robot is performing incorrectly? When the behavior of the robot displays an inadequate attempt, the instructor should redirect or educate attention by repeating or modifying an action. The catch is that the human's degrees of freedom might not map at all well onto those of the robot (for example, see Bicchi (2000) for the variety of forms taken by the dexterous robot's hands). In some ways, though, the difference between child and robot is not so great. Even though

[3] In forward models, information from both the body and the environment inform emerging action.

the adult human shares general bodily architecture with the child, the adult cannot share or access the child's neural processes nor those of the robot. (As is evident in these passages a challenge, only partially met, is to unpackage aspects of ecological realist and computational approaches that do not at first converge.)

From a computational perspective, the success of training results from the child's ability to develop a forward model that can link the (unobservable) internal processes to their observable consequences. The solution must relate to developing a means to discover the goals that characterize the success of diverse actions. Given that a robot can receive symbolic messages, telling the robot how to link what is observed to the knowledge with which the robot has been endowed might facilitate finding a better solution. Such messages would make explicit what caregivers must do by ostension during assisted imitation of human infants. A whole other issue, of course, for practical robots is what happens when the "caregiver" is another robot. When does an "adult" robot need to train a "young" robot rather than simply downloading the program and parameters that encode the fruits of its own experience?

14.5 Discussion

These "naturalistic investigations" delineated some of the ways in which imitation, especially assisted imitation, might contribute to communicative development. Engaging in action sequences with the caregiver cultivated a precursor to language: negotiating a common understanding of ongoing events. Implications from this research may clarify documented differences in the communicative and didactic behaviors of non-human and human primates, and inform studies of early lexical development, as well as the transition from single-element communications to two-word utterances.

14.5.1 Differentiating the difference between non-human primates and humans

The empirical literature documents that monkeys imitate each other very infrequently, if at all (Whiten and Ham, 1992; Bard and Russell, 1999; Visalberghi and Fragaszy, 2002). Chimpanzees imitate some actions of others in the wild (Quiatt and Itani, 1994; Greenfield *et al.* 2000), imitate the sequential and hierarchical structure of action in experimental, laboratory settings (Whiten, 1992), but learn a broader range of complex actions with objects when raised by humans (Savage-Rumbaugh, 1986; Tomasello *et al.*, 1993; Savage-Rumbaugh and Lewin, 1994; Tomasello and Call, 1997). However, the pace and extent of their imitation is very limited with respect to that of humans (see Stanford (this volume) for a critique of this literature; Greenfield (this volume) for a more expansive view of the commonalities between monkeys, apes, and humans). Indeed, the vast majority of human children do imitate, albeit to varying degrees at different ages and for behaviors that differ in modality and complexity of content (Eckerman, 1993; Nadel and Butterworth, 1999).

Why does teaching occur so massively among humans, whereas very sparse evidence documents that chimpanzees or any other non-human primates do so? What basic abilities allow humans to tutor their prelinguistic infants that remain out of reach or, perhaps, out of sight for non-human primates? Human caregivers invite their offspring to imitate new behaviors, and intervene and prompt them when their attempts go awry. To teach effectively, a caregiver must know how to engage in a proffered activity and see whether the novice can use similar methods to achieve such an aim. Further, as infants respond to caregiver guidance, their embodied misunderstandings of caregiver messages often inform caregivers' subsequent feedback. Detecting what infants cannot do and the affordances they miss perceiving is inextricably linked to effectively educating infants' attention, so that their subsequent attempts fall closer to the mark.

Some have noted that humans learn most affordances in a social/cultural context (Heft, 1989), but have had little to say about that process, except to stress that practice enhances perceptual differentiation (E. Gibson, 2002). Gibson and Rader (1979) reasoned that infants and young children pick up information not relevant to the task at hand as they simply may not know what to notice. In the same vein, Sharrock and Coulter (1998) noted that animals and prelinguistic infants distinguish one thing from another, but they do not know their significance within cultural activities. This present study of assisted imitation documents that caregivers actively, not indirectly, fill in this gap day in and day out. During the most humdrum activities they educate attention to what just such a body can do in concert with the objects on hand to achieve a specific aim while situated in a particular environment.

Some scholars have suggested that Theory of Mind abilities serve as a basis for teaching/tutoring (Meltzoff and Moore, 1995, 1999; Bråten, 2002). as well as language evolution and development (see Gordon (this volume) for discussion of the relation between the evolution of Theory of Mind and of language). In this view, the expert can *mind read*, inferring what the novice knows and does not know. Reboul (2004), however, argues that precursors of Theory of Mind, rather than full-blown abilities, may be sufficient to engage in complex social–cultural activities. That is, detecting the visual appearance of body and face, direction of gaze, and the target or aim of another's action may contribute to *behavior reading*, understanding another's actions (Sterelny, 2002). Numerous recent investigations of non-human primate perception document that chimpanzees engage in behavior reading. That is, they can engage in joint attention and observational learning (Greenfield *et al*., 2000) and understand that if another faces or orients toward a particular configuration of objects and/or creatures, the other may act in particular ways (Boesch, 1993; Povinelli *et al*., 2000; Sterelny, 2002). According to Hare *et al*. (2001), chimpanzees can prospectively perceive differences in the aims of others' actions. For example, if the subordinate chimpanzee notices that the more dominant one cannot see something desirable, the less dominant often will take the object or food quickly before the other notices. Finally, bonobos can orient others with a declarative hand point to an individual who is out of their line of sight (author's field notes: Kanzi, July 1987). In all these examples, I suggest that chimpanzees *perceive*

what the other perceives; they watch *what* others do and can "see" *what* another is likely to do next. However, do such achievements, in turn, provide enough information or display the abilities necessary for teaching new actions with objects to a novice? Apparently not, as non-human primate infants rarely receive "active" teaching (for two recorded examples, see Boesch, 1993), and to date there is no evidence for "hands on" guidance (embodying of another) while manipulating objects. Non-human primates who observe and sometimes imitate new activities, albeit not at the rate of humans nor in quite the same way, take many years to learn skills, such as tool use (Boesch, 1993; K. R. Gibson, 1993). In contrast, human infants and enculturated chimpanzees who benefit from the guidance of caregivers learn similar skills quite rapidly (Savage-Rumbaugh, 1994).

To teach or tutor, experts must *perceive how others perceive*. More competent individuals do not simply perceive *what* others do or do not perceive. Teachers observe the public and emergent process of just *how* infants' knowing/learning what the body is and does educates perceiving affordances.[4] As action unfolds, the infants' misdoings, such as misaligning hand and object, misorienting of one object to another, deviations in path of action, and so on, pinpoint possible misperceiving of affordances as well as lack of bodily skill. That is, infants' embodied misunderstandings of caregiver messages display both *what* and *how* they "miss" perceiving as more adept members do. Thus, human caregivers can and do detect the (culturally) unskilled behavior of novices which frequently informs their subsequent feedback. At this point, caregivers often embody infants, so they can help them get and keep in touch with *how* the body and the objects of action work together. In addition, as the interaction unfolds dynamically, the infant–novice can learn how to "look" at the ways something can be seen/perceived in order to achieve some aim. For example, an infant first sees and feels in arm and hand the resistance of the protrusion on the pop bead before just the right amount of force at a particular angle pushes it through the opening on another bead. Without that perceptual information, the infant does not have the embodied understanding of what body and beads do together in and over time to get them to concatenate. Caregiver assistance may be the source that speeds up learning and increases skill by educating *what and how* the infant perceives as body and environment meet.

To get things done, members continuously constitute a joint understanding of what is happening.[5] These interactions also may be central to communicative development. As a by-product of the constant monitoring and engagement in interaction during mundane daily activities, these practices may pave the way to early word learning. In particular, engaging in these activities may provide the means to grasp important prerequisites that underlie communicating with language. These basics include knowing that words have an instrumental effect on the receiver of a message (Braunwald, 1978, 1979), words refer (Bates *et al*., 1979; Schlesinger, 1982; Zukow-Goldring, 1997;

[4] People can perceive with accuracy where another is looking (J. J. Gibson and Pick, 1963) as well as detect affordances for others (Stoffregen *et al*., 1999).

[5] See Goodwin (2000) for a recent edited collection of studies from linguistic anthropology that explores how people make the unfolding structure and organization of events visible and meaningful to one another.

Zukow-Goldring and Rader, 2001), and coparticipants share or negotiate a common understanding of ongoing events (Moerman, 1988; Zukow, 1990; Macbeth, 1994; Zukow-Goldring, 1997).

14.5.2 Perceiving reference: a key to early lexical development

How do infants unlock the puzzle of reference, the mundane fact that words refer to objects, actions, attributes, and more? That is, how do infants learn the relation between the stream of speech and the unceasing flow of events? While this sounds like a simple question, its answer remains elusive.

The preponderance of research investigating lexical development has tested hypotheses proposing that innate predispositions (Markman, 1989), cognitive prerequisites (Clark, 1993; Golinkoff *et al*., 1994), and/or social abilities (Tomasello, 1988, 2001) underlie lexical development, usually from 15 months onward. Some 6 months earlier, however, infants begin to display comprehension and production of their first words. A large body of work documents which types of words and their tokens emerge when and in what order (Fenson *et al*., 1994). By and large, this literature assesses what the infant already knows, not how the infant comes to know. Caregivers talk about what they are doing as they do it. Given these circumstances, how is consensus achieved as the child becomes an adept member of the community? My approach integrates discovering the interactional methods or practices that inform perceptual differentiation, the assembling of action sequences, and the detecting of word meaning, despite the fact that many studies of language acquisition assume that gestures entail ambiguity of reference (Schlesinger, 1982; Markman, 1989). These authors rely on Quine's classic essay (1960) in which he discussed the ambiguity of reference entailed in, say, speaking about and pointing to a rabbit. But caregivers tend to focus attention with precision. They do not simply say an unfamiliar word (such as Quine's *gavagai*) while pointing. Instead, caregivers may rub a rabbit's fur while saying "fur"; trace the topography of its ears while saying "ear," stroke the entire rabbit or rotate the whole animal when saying "rabbit," etc. (Zukow 1990; Reed, 1995; Zukow-Goldring, 1996). Successful teaching entails marking the correspondence between what is said and what is happening.

Zukow-Goldring and colleagues (Zukow, 1990; Zukow-Goldring, 1997; Zukow-Goldring and Rader, 2001) derived hypotheses from naturalistic, longitudinal studies of early word learning, in Mexico and the United States that have been confirmed by Gogate and colleagues (Gogate *et al*., 2000, 2001). Zukow-Goldring has shown how caregivers bracket ongoing actions with gestures that direct the child's attention to perceptual information embodied in action sequences as well as the perceivable correspondence (in offset/onset, tempo, and rhythm) between word and referent. Attention to movement and synchrony in gesture and speech facilitates infants' detection of the correspondence between two fundamentally different kinds of things: words and aspects of ongoing events. Caregivers nest their messages in higher-order amodal regularities with a temporal basis that make word and referent "stand in" relation to each other and

"stand out" from other possibilities. For example, a caregiver may "say-and-do" synchronously by saying *sti::cky*, in a rough, raspy voice while simultaneously displaying the irregular surface below her fingers as she runs them over uneven broom bristles, assisting the infant to "see-and-hear" the correspondence between the two. Understanding infants' perceptual abilities and caregiver practices underlying the emergence of early word comprehension is key, because language development without knowing what and that words mean cannot proceed.

Enlarging on the import of these practices, Falk (2004) has explored the possibility that the multi-modal characteristics of contemporary mother–infant gesture and vocalizations in both chimpanzees and humans may have derived from the behavior of early hominins. He hypothesized that early communicative behaviors arose in the form of a "primal" song. That is, when foraging for food, early hominins may have kept in touch with their infants in a non-tactile way by using a melodic, comforting vocalizations as a sort of vocal rocking or soothing after putting them down. Elaborating from these early communicative behaviors, he speculated that these messages might have the power to explain the evolution of protolanguage that would eventually lead to conventionalized meanings and language itself.

Returning to present-day mothers of human infants, some indirect evidence suggests that assisted imitation may also cultivate production of early words and the comprehension of caregiver messages. The longitudinal data from both English- and Spanish-speaking families confirm that caregiver messages communicate perceptual structure and/or semantic functions that are a step or two ahead of those expressed in infants' speech (Zukow-Goldring, 1997). These findings document a tie between caregiver input and subsequent production of words. Further, caregivers of less advanced infants (not necessarily younger infants) other-regulate their infants' attention by *embodying* and *show*ing most frequently, shifting to *demonstrations, points*, and eventually *looks* as the infants gradually self-regulate their own attention. When caregivers provide more perceptual structure after initial misunderstandings of caregiver messages, infants do comprehend what is meant significantly more often (Zukow-Goldring, 1996, 2001). A challenge for the future is to determine if there is a relation between this set of caregiver gestures and infant gestural production. (For other aspects of learning deictic and symbolic gestures, see Capirci *et al*., 1996; Iverson *et al*., 1994.) However, *embodying* would likely appear last in the infant's repertoire rather than first, as the embodier must detect *how* the other might misperceive and then redirect attention to a more informative action–perception coupling. In contrast, pointing to *what* you want is not as complex a task.

14.5.3 Communicative action and gesture precede and inform two-word utterances

Greenfield and Smith (1976) emphasized that language development arises from and builds on the infants' non-verbal understanding of events, their "world of experience," a world embedded in social interaction (see also Greenfield, this volume). That world is

"*in* culture"; it is the only world the infant could ever know (Costall, 1989, p.19; Zukow-Goldring, 1997). This section delineates some of the links between early action sequences and two-word utterances, the continuity between the content and structure of these messages as the infant develops, and the role of dialogue during social interaction.

Some infant actions, although unintentionally communicative, do have a perlocutionary force or effect on the caregiver, when noticed. If a caregiver notices an infant pointing for the self to, say, a toy dog, the caregiver will often orient to the toy and say, *dog*, although the infant has not engaged in pointing for the caregiver's benefit. On the other hand, an infant who reached and whined for a piece of *bolillo* (roll) remained empty-handed when he did not catch his aunt's gaze. In contrast, 6 weeks later, under the same interactional and situational circumstances, he waited to catch and alternate gaze to the desired food. This time he very quickly received some roll (author's field notes, 1981–2). Thus, beginning in the months just prior to the emergence of speech, infants request action and/or objects by reaching and whining while alternating gaze or by placing objects in the caregiver's hands for manipulation: *proto-imperatives*. They also direct the attention of others by pointing to elements in the environment while alternating gaze with caregivers: *proto-declaratives* (Bates *et al*., 1979). Most infants begin to produce such communicative actions between 8 and 10 months of age (Bates, 1976; Volterra *et al*., 2005). Other single-element messages comprised of a symbolic gesture or a one-word utterance emerge somewhat later, nearer to the end of the first year of life or early in the second year (Acredolo and Goodwyn, 1985; Butcher and Goldin-Meadow, 2000; Volterra *et al*., 2005).[6] Single-element action or gestural communications predominate until around 16 months of age; thereafter words gradually become more numerous (Iverson *et al*., 1994).

During the single-element period, infant communications express semantic functions, such as agents, actions/states, objects, recipients of action, locations, instruments, and more (Greenfield and Smith, 1976; Zukow-Goldring, 1997). In the interim between single-element messages and two-word utterances, infants begin to combine communicative actions, symbolic gestures, and single words in a variety of ways from about 17–18 months of age (Greenfield and Smith, 1976; Greenfield *et al*., 1985; Butcher and Goldin-Meadow, 2000; Volterra *et al*., 2005). The messages may convey *redundant* (semantic) information by referring to the same referent (waving bye-bye while saying *ciao), complementary* information by enacting a deictic[7] gesture directing attention to a referent (pointing to a cup, while saying *cup*), and *supplementary* or *syntagmatic*[8] information by expressing different semantic functions that add information to each other (pointing to a record player, while saying *on*) (Greenfield and Smith, 1976; Greenfield

[6] Note that definitions of gesture vary from one researcher to another. For instance, Volterra and her colleagues (2005) include communicative actions in which the infant is touching something or someone, whereas Butcher and Goldin-Meadow (2000) exclude such actions.

[7] *Deictic* refers to a word or gesture that depends on the context of its production for its meaning, e.g., *I* refers to the speaker; the trajectory of a point intersects with a particular target of attention.

[8] Butcher and Goldin-Meadow (2000) use the term *complementary* instead of *supplementary* or *syntagmatic* to designate such combinations, e.g., pointing to a cup and saying *mine*.

et al., 1985; Capirci *et al*., 1996). In addition, the infant may produce successive single-word utterances first (*Bib. Off.)*, then two-element messages combining gesture and word (reaching and whining for a banana, while saying *nana)*, and eventually multi-word utterances (*eat nana)*, suggesting that these transitional forms may underpin syntactic expression (Greenfield and Smith, 1976; Greenfield *et al*., 1985).

Both Greenfield and Volterra and their colleagues explicitly elaborate some potential links between two-element *supplementary* messages and the emergence of syntax. Further, both research groups have explored the ways that dialogue with the caregiver nurtures these communicative developments. Multi-word caregiver messages and question–answer sequences about a particular activity may set the stage for subsequent messages on the "same" topic built from contributions of both the caregiver and the infant. These jointly expressed messages conveying different aspects of the ongoing interaction may precede and promote two-element messages that the infant later will produce autonomously.

Butcher and Goldin-Meadow's research (2000) documents that the onset of combinations in which gesture and speech expressed different information correlated significantly with the onset of two-word utterances. The ability to concatenate two elements conveying different aspects of a situation within a single communicative act appears necessary but not sufficient to guarantee two-word speech as no infant in their study produced two-word speech without first expressing such gesture–word combinations. During the latter part of the second year, nearly all infants had made the transition to multi-word utterances.

This body of research highlights the continuity between prelinguistic and linguistic communication, underscoring the crucial role of action sequences embedded in social interaction and dialogue to the emergence of gestural communication and ultimately for the emergence of language.

14.5.4 Summary

Caregivers both direct attention to aspects of the ongoing events and tutor actions to "achieve consensus." In a sense, they bring the child to share prospectively with the caregiver how someone may discover which effectivities can "deliver" the affordances offered by objects in the environment. These interactional opportunities give infants crucial practice in (and a refining of) what to notice and do, and when to do it.

In the context of understanding how the methods of the caregiver correspond to the expanding capabilities of the child, note that often developmental researchers and scholars study affective, motor, perceptual, and cognitive development separately. Caregivers do not. During the prelinguistic and one-word periods, caregivers prepare infants to imitate by assisting them "to see what to do" before they can "do what they see" others doing. Day in and day out, they cultivate imitation within mundane daily activities with gestures. They animate and direct their infants' attention to their own and others' bodily movements as well as making prominent what the environment offers for action.

In the process of learning a new skill, especially when embodied or put through the motions by the caregiver, infants directly experience that the self is "like the other." Thus, embedded in grasping the *transfer of skill* problem are opportunities to see and feel solutions to the *correspondence* problem, detecting the match between self and other. Humans who eventually learn/understand that the self is "like the other" cultivate abilities in their young that contribute to imitating, tutoring, and communicating. The mirror system offers a means to clarify in what manner human and non-human primates as well as intelligent automata understand what they see other intelligent agents doing, what abilities and perceptual information underlie learning to do what they see others do, and much more.

Acknowledgments

The Spencer Foundation and the Michelle F. Elkind Foundation provided funds for this research. I wish to express my gratitude to colleagues who generously commented on prior versions of this manuscript: Harry Heft, Thomas Stoffregen, Dankert Vedeler, Cathy Dent-Read, Alan Costall, Blandine Bril, Michael Arbib (our indefatigable editor), and three anonymous reviewers. I extend my thanks to the caregivers and infants who participated in these investigations as well as Sarah Anderson and Heewoon Chung for their assistance.

References

Acredolo, L. P., and Goodwyn, S. W., 1985. Symbolic gesture in language development: a case study. *Hum. Devel.* **28**: 40–49.

Arbib, M. A., 2002. The Mirror System, imitation, and evolution of language. In C. L. Nehaniv and K. Dautenhahn (eds.) *Imitation in Animals and Artifacts*. Cambridge MA: MIT Press, pp. 229–280.

Atkeson, C. G., and Schaal, S., 1997. Robot learning from demonstration. *Int. Conference on Machine Learning*, San Francisco, CA, pp. 12–20.

Bahrick, L. E., and Pickens, J. N., 1994. Amodal relations: the basis for intermodal perception and learning in infancy. In D. J. Lewkowicz and R. Lickliter (eds.) *The Development of Intersensory Perception: Comparative Perspectives*. Hillsdale, NJ: Lawrence Erlbaum, pp. 205–233.

Bard, K. A., and Russell, C. L., 1999. Evolutionary foundations of imitation: social, cognitive and developmental aspects of imitative processes in non-human primates. In J. Nadel and G. Butterworth (eds.) *Imitation in Infancy*. Cambridge, UK: Cambridge University Press, pp. 89–123.

Bates, E., 1976. *Language and Context: The Acquisition of Pragmatics*. New York: Academic Press.

Bates, E., Begnini., L., Camaioni, L., Bretherton, I., and Volterra, V., 1979. *The Emergence of Symbols: Cognition and Communication in Infancy*. New York: Academic Press.

Bicchi, A., 2000. Hands for dexterons manipulation and robust grasping: a difficult road towards simplicity. *IEEE Trans. Robot. Automat.* **16**: 652–662.

Billard, A., 2002. Imitation: a means to enhance learning of a synthetic proto-language in an autonomous robot. In C. L. Nehaniv and K. Dautenhahn (eds.) *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, pp. 281–310.

Boesch, C., 1993. Aspects of transmission of tool-use in wild chimpanzees. In K. R. Gibson and T. Ingold (eds.) *Tools, Language, Cognition in Human Evolution*. Cambridge, UK: Cambridge University Press, pp. 171–183.

Bråten, S., 2002. Altercentric perception by infants and adults in dialogue. In M. I. Stamenov and V. Gallese (eds.) *Mirror Neurons and the Evolution of Brain and Language*. Amsterdam, Netherlands: Benjamin, pp. 273–294.

Braunwald, S., 1978. Context, word and meaning: toward a communicational analysis of lexical acquisition. In A. Lock (ed.) *Action, Gesture, and Symbol: The Emergence of Language*. London: Academic Press, pp. 285–327.

1979. On being understood: the listener's contribution to the toddler's ability to communicate. In P. French (ed.) *The Development of Meaning*. Hiroshima, Japan: Bunko Hyonron Press, pp. 71–113.

Breazeal, C., and Scassellati, B., 2002. Robots that imitate humans. *Trends Cogn. Sci.* **6**: 481–487.

Bril, B., 2004. Learning to use scissors: adult's scaffolding and task properties. *Proceedings Int. Congress of Psychology*, Beijing, China, pp. 545–546.

Bruner, J., 1983. *Children's Talk: Learning to Use Language*. New York: Norton.

Butcher, C., and Goldin-Meadow, S., 2000. Gesture and the transition from one- to two-word speech: when hand and mouth come together. In D. McNeill (ed.) *Language and Gesture*. Cambridge, UK: Cambridge University Press, pp. 235–257.

Byrne, R. W., 2003. Imitation as behaviour parsing. *Phil. Trans. Roy. Soc. London B* **358**: 529–536.

Capirci, O., Iverson, J. M., Pizzuto, E., and Volterra, V., 1996. Gestures and words during the transition to two-word speech. *J. Child Lang.* **23**: 645–673.

Clark, E. V., 1993. *The Lexicon in Acquisition*. Cambridge, UK: Cambridge University Press.

Costall, A., 1989. A closer look at "direct perception." In A. Gellatly, D. Rodgers, and J. A. Sloboda (eds.) *Cognition and Social Worlds.* Oxford, UK: Clarendon Press, pp. 10–21.

1995. Socializing affordances. *Theory Psychol.* **5**: 1–27.

Dautenhahn, K., and Nehaniv, C. L., 2002. The agent-based perspective on imitation. In C. L. Nehaniv and K. Dautenhahn (eds.) *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, pp. 1–40.

Eckerman, C. O., 1993. Toddlers' achievement of coordinated action with conspecifics: a dynamic systems perspective. In L. B. Smith and E. Thelen (eds.) *A Dynamic Systems Approach to Development*. Cambridge, MA: MIT Press, pp. 333–357.

Falk, D., 2004. Prelinguistic evolution in early hominins: whence motherese? *Behav. Brain Sci.* **27**: 491–503.

Fenson, L., Dale, P. S., Reznick, J. S., *et al.*, 1994. Variability in early communicative development. *Monogr. Soc. Res. Child Devel.* **59**.

Gibson, E. J., 1969. *Principles of Perceptual Learning and Development*. New York: Appleton-Century-Crofts.

2002. *Perceiving the Affordances: A Portrait of Two Psychologists*. Mahwah, NJ: Lawrence Erlbaum.

Gibson, E. J., and Pick, A. D., 2000. *An Ecological Approach to Perceptual Learning and Development*. New York: Oxford University Press.
Gibson, E. J., and Rader, N., 1979. Attention: the perceiver as performer. In G. A. Hale and M. Lewis (eds.) *Attention and Cognitive Development*. New York: Plenum Press, pp. 1–21.
Gibson, J. J., 1979. *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
Gibson, J. J., and Pick, A. D., 1963. Perception of another person's looking behavior. *Am. J. Psychol*. **76**: 386–394.
Gibson, K. R., 1993. Introduction: generative interplay between technical capacities, social relations, imitation and cognition. In K. R. Gibson and T. Ingold (eds.) *Tools, Language, Cognition in Human Evolution*. Cambridge, UK: Cambridge University Press, pp. 131–137.
Gogate, L. J., Bahrick, L. E., and Watson, J. D., 2000. A study of multimodal motherese: the role of temporal synchrony between verbal labels and gestures. *Child Devel*. **71**: 878–894.
Gogate, L. J., Walker-Andrews, A. S., and Bahrick, L. E., 2001. The intersensory origins of word comprehension: an ecological–dynamic systems view. *Devel. Sci*. **4**: 1–18.
Golinkoff, R., Mervis, C., and Hirsh-Pasek, K., 1994. Early object labels: the case for a developmental lexical principles framework. *J. Child Lang*. **21**: 125–155.
Goodwin, C. (ed.), 2000. Vision and inscription in practice. *Mind, Culture, and Activity* **7**.
Greenfield, P. M., 1972. Cross-cultural studies of mother–infant interaction: toward a structural–functional approach. *Hum. Devel.* **15**: 131–138.
Greenfield, P. M., and Smith, J., 1976. *The Structure of Communication in Early Language Development*. New York: Academic Press.
Greenfield, P. M., Reilly, J., Leaper, C., and Baker, N., 1985. The structural and functional status of single-word utterances. In M. D. Barrett (ed.) *Children's Single-Word Speech*. Chichester, UK: John Wiley, pp. 233–267.
Greenfield, P. M., Maynard, A. E., Boehm, C., and Schmidtling, E. Y., 2000. Cultural apprenticeship and cultural change: tool learning and imitation in chimpanzees and humans. In S. T. Parker, J. Langer, and M. L. McKinney (eds.) *Biology, Brains and Behaviour*. Santa Fe, NM: SAR Press, pp. 237–277.
Hare, B., Call, J., and Tomasello, M., 2001. Do chimpanzees know what conspecifics know? *Anim. Behav*. **61**: 139–151.
Heft, H., 1989. Affordances and the body: an intentional analysis of Gibson's ecological approach to visual perception. *J. Theory Soc. Behav*. **19**: 1–30.
Iacoboni, M., Woods, R. P., Brass, M., *et al*., 1999. Cortical mechanisms of human imitation. *Science* **286**: 2526–2528.
Ingold, T., 2000. *The Perception of the Environment: Essays in Livelihood, Dwelling and Skill*. London: Routledge.
Iverson, J. M., Capirci, O., and Caselli, M. C., 1994. From communication to language in two modalities. *Cogn. Devel*. **9**: 23–43.
Jordan M. I., and D. E. Rumelhart, 1992. Forward models: supervised learning with a distal teacher. *Cogn. Sci*. **16**: 307–354.
Kadar, E., and Shaw, R. E., 2000. Toward an ecological field theory of perceptual control of locomotion. *Ecol. Psychol*. **12**: 141–180.
Kendon, A., 2004. *Gesture: Visible Action as Utterance*. Cambridge, UK: Cambridge University Press.

Macbeth, D., 1994. Classroom encounters with the unspeakable: "Do you see, Danelle?" *Discourse Proc.* **17**: 311–335.
Mace, W., 1977. Ask not what's in your head, but what your head's inside of. In R. E. Shaw and J. Bransford (eds.) *Perceiving, Acting, and Knowing*. Hillsdale, NJ: Lawrence Erlbaum, pp. 43–65.
Markman, E. M., 1989. *Categorization and Naming in Children: Problems of Induction*. Cambridge, MA: MIT Press.
Martini, M., and Kirkpatrick, J., 1992. Parenting in Polynesia: a view from the Marquesas. In J. L. Roopnarine and D. B. Carter (eds.) *Parent–Child Socialization in Diverse Cultures*. Norwood, NJ: Ablex, pp. 199–223.
Maynard, A. E., 2002. Cultural teaching: the development of teaching skills in Zinacantec Maya sibling interactions. *Child Devel.* **73**: 969–982.
McGovern, A., and Barto, A. G., 2001. Automatic discovery of subgoals in reinforcement learning using diverse density. *Proceedings 18th Int. Conference on Machine Learning*, San Francisco, CA, pp. 361–368.
Meltzoff, A. N., and Moore, M. K., 1995. Infants' understanding of people and things: from body imitation to folk psychology. In J. Bermúdez, A. J. Marcel, and N. Eilan (eds.) *The Body and the Self*. Cambridge, MA: MIT Press, pp. 43–69.
1999. Persons and representation: why infant imitation is important for theories of human development. In J. Nadel and G. Butterworth (eds.) *Imitation in Infancy*. Cambridge, UK: Cambridge University Press, pp. 9–35.
Michaels, C. F., and Carello, C., 1981. *Direct Perception*. Englewood Cliffs, NJ: Prentice-Hall.
Moerman, M., 1988. *Talking Culture: Ethnography and Conversation Analysis*. Philadelphia, PA: University of Pennsylvania Press.
Moore, C., 2005. *Discussion: Imitation, Identification, and Self–Other Awareness*. Atlanta, GA: Society for Research on Child Development.
Moore, C., and Dunham, P. (eds), 1995. *Joint Attention: Its Origins and Role in Development*. Hillsdale, NJ: Lawrence Erlbaum.
Nadel, J., and Butterworth, G. (eds.), 1999. *Imitation in Infancy*. Cambridge, UK: Cambridge University Press.
Nadel, J., Guérini, C., Pezé, A., and Rivet, C., 1999. The evolving nature of imitation as a format for communication. In J. Nadel and G. Butterworth (eds.) *Imitation in Infancy*. Cambridge, UK: Cambridge University Press, pp. 209–234.
Nehaniv, C. L., and Dautenhahn, K., 2002. The correspondence problem. In K. Dautenhahn and C. L. Nehaniv (eds.) *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, pp. 41–61.
Oztop, E., and Arbib, M. A., 2002. Schema design and implementation of the grasp-related mirror neuron system. *Biol. Cybernet.* **87**: 116–140.
Oztop, E., Bradley, N. S., and Arbib, M. A., 2004. Infant grasp learning: a computational model. *Exp. Brain Res.* **158**: 480–503.
Piaget, J., 1962. *Play, Dreams, and Imitation in Childhood*. New York: Norton.
Povinelli, D. J., Bering, J. M., and Giambrone, S., 2000. Toward a science of other minds: escaping the argument by analogy. *Cogn. Sci.* **24**: 509–541.
Quiatt, D., and Itani, J., 1994. *Hominid Culture in Primate Perspective*. Niwot, CO: University Press of Colorado.
Quine, W. V. O., 1960. *Word and Object*. New York: John Wiley.
Reboul, A., 2004. Evolution of language from theory of mind or coevolution of language and theory of mind? Available at http://www.interdisciplines

Reed, E. S., 1995. The ecological approach to language development. *Lang. Communi.* **15**: 1–29.
Rizzolatti, G., and Arbib, M. A., 1998. Language within our grasp. *Trends Neurosci.* **21**: 188–194.
Savage-Rumbaugh, E. S., 1986. *Ape Language: From Conditioned Response to Symbol*. New York: Columbia University Press.
Savage-Rumbaugh, E. S., and Lewin, R., 1994. *Kanzi: The Ape at the Brink of the Human Mind*. New York: John Wiley.
Schaall, S., 1999. Is imitation learning the route to humanoid robots? *Trends Cogn. Sci.* **3**: 233–242.
Schlesinger, I. M., 1982. *Steps to Language: Toward a Theory of Native Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum.
Schutz, A., 1962. *Collected Papers*, vol. 1, *The Problem of Social Reality*. Leiden, Netherlands: Martinus Nijhoff.
Sharrock, W., and Coulter, J., 1998. On what we can see. *Theory Psychol.* **8**: 147–164.
Shaw, R., 2001. Processes, acts, and experiences: three stances on the problem of intentionality. *Ecol. Psychol.* **13**: 275–314.
Shaw, R., and Turvey, M., 1981. Coalitions as models of ecosystems: a realist perspective on perceptual organization. In M. Kubovy and J. R. Pomerantz (eds.) *Perceptual Organization*. Hillsdale, NJ: Lawrence Erlbaum, pp. 343–415.
Snow, C. E., Perlmann, R., and Nathan, D., 1987. Why routines are different: toward a multiple-factor model of the relation between input and language acquisition. In K. E. Nelson and A. van Kleeck (eds.) *Children's Language*, vol. 6. Hillsdale, NJ: Lawrence Erlbaum, pp. 65–97.
Spelke, E. S., 1979. Perceiving bimodally specified events in infancy. *Devel. Psychol.* **15**: 626–636.
Sterelny, K., 2002. Primate worlds. In C. Heyes and L. Huber (eds.) *The Evolution of Cognition*. Cambridge, MA: MIT Press, pp. 143–162.
Stoffregen, T. A., 2003. Affordances as properties of the animal–environment system. *Ecol. Psychol.* **15**: 115–134.
Stoffregen, T., and Bardy, B., 2001. On specification and the senses. *Behav. Brain Sci.* **24**: 195–261.
Stoffregen, T. A., Gorday, K. M., Sheng, Y., and Flynn, S. B., 1999. Perceiving affordances for another person's actions. *J. Exp. Psychol. Hum. Percept. Perform.* **25**: 120–136.
Thelen, E., and Smith, L., 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.
Thorndike, E. L., 1898. Animal intelligence: an experimental study of the associative processes in animals. *Psychol. Rev. Monogr.* **2**.
Tomasello, M., 1988. The role of joint attentional processes in early language development. *Lang. Sci.* **10**: 69–88.
2001. Perceiving intentions and learning words in the second year of life. In M. Bowerman and S. C. Levinson (eds.) *Language Acquisition and Conceptual Development*. Cambridge, UK: Cambridge University Press, pp. 132–158.
Tomasello, M., and Call, J., 1997. *Primate Cognition*. New York: Oxford University Press.
Tomasello, M., Kruger, A. C., and Ratner, H. H., 1993. Cultural learning. *Behav. Brain Sci.* **16**: 495–552.

Tomasello, M., Savage-Rumbaugh, S., and Kruger, A. C., 1993. Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Devel.* **64**: 1688–1705.

Turvey, M., 1992. Affordances and prospective control: an outline of the ontology. *Ecol. Psychol.* **4**: 173–187.

Turvey, M., Shaw, R., Reed, E., and Mace, W., 1981. Ecological laws of perceiving and acting. *Cognition* **9**: 237–304.

Uzgiris, I., 1991. The social context of imitation. In M. Lewis and S. Feinman (eds.) *Social Influences and Socialization in Infancy*. New York: Plenum Press, pp. 215–251.

Visalberghi, E., and Fragaszy, D., 2002. "Do monkeys ape?" – ten years later. In K. Dautenhahn and C. L. Nehaniv (eds.) *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, pp. 471–499.

Volterra, V., Caselli, M. C., Capirci, O., and Pizzuto, E., 2005. Gesture and the emergence and development of language. In M. Tomasello and D. Slobin (eds.) *Beyond Nature–Nurture: Essays in Honor of Elizabeth Bates*. Mahwah, NJ: Lawrence Erlbaum, pp. 3–40.

Vygotsky, L. S., 1978. *Mind in Society*. (M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, eds.) Cambridge, MA: Harvard University Press.

Watson-Gegeo, K. A., and Gegeo, D. W., 1986. Calling-out and repeating routines in Kwara'ae children's language socialization. In B. Schieffelin, and E. Ochs (eds.) *Language Socialization across Cultures*. New York: Cambridge University Press, pp. 17–50.

1999. (Re)modeling culture in Kwara'ae: the role of discourse in children's cognitive development. *Discourse Stud.* **1**: 241–260.

Wertsch, J. V., McNamee, G. D., McLane, J. B., and Budwig, N., 1980. The adult–child dyad as a problem-solving system. *Child Devel.* **51**: 1215–1221.

Whiten, A., 2002. Imitation of sequential and hierarchical structure in action: experimental studies with children and chimpanzees. In K. Dautenhahn and C. L. Nehaniv (eds.) *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, pp. 192–209.

Whiten, A., and Ham, R., 1992. On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. In P. J. B. Slater, J. S. Rosenblatt, C. Beer, and M. Milinski (eds.) *Advances in the Study of Behavior*, vol. 21. San Diego, CA: Academic Press, pp. 239–283.

Witherington, D., 2005. A comparison of two conceptualizations for the concept of affordance. *Proceedings Int. Conference on Perception and Action*, Asilomar, CA.

Wood, D., Bruner, J. S., and Ross, G., 1976. The role of tutoring in problem solving. *J. Child Psychol. Psychiatr.* **17**: 89–100.

Zukow, P. G., 1989. Siblings as effective socializing agents: evidence from Central Mexico. In P. G. Zukow (ed.) *Sibling Interactions across Cultures: Theoretical and Methodological Issues*. New York: Springer-Verlag, pp. 79–105.

1990. Socio-perceptual bases for the emergence of language: an alternative to innatist approaches. *Devel. Psychobiol.* **23**: 705–726.

1996. Sensitive caregivers foster the comprehension of speech: when gestures speak louder than words. *Early Devel. Parent.* **5**: 195–211.

1997. A social ecological realist approach to the emergence of the lexicon: educating attention to amodal invariants in gesture and speech. In C. Dent-Read and P. Zukow-Goldring (eds.) *Evolving Explanations of Development: Ecological*

Approaches to Organism-Environment Systems. Washington, DC: American Psychological Association, pp. 199–250.

2001. Perceiving referring actions: Latino and Euro-American infants and caregivers comprehending speech. In K. E. Nelson, A. Aksu-Koc, and C. Johnson (eds.) *Children's Language*, vol. 11. Hillsdale NJ: Lawrence Erlbaum, pp. 139–163.

Zukow-Goldring, P., and Arbib, M. A., in press. Affordances, effectivities, and assisted imitation: caregivers and the direction of attention. *Neurocomp*.

Zukow-Goldring, P., and Rader, N., 2001. Perceiving referring actions: a commentary on Gogate, Walker-Andrews, and Bahrick. *Devel. Sci*. **4**: 28–30.