

5 Perceiving intentions and learning words in the second year of life

Michael Tomasello

Emory University

Most studies of the cognitive bases of language acquisition are concerned with semantic content: how young children conceptualize the referents of particular linguistic items and expressions. But being able to conceive of all the possible referents another person might intend to indicate is not enough. To acquire the conventional use of a new linguistic item a learner must also be able to identify which of these conceivable referents another person is attempting to single out when using that item in a particular communicative circumstance. This raises issues of *social* cognition and how young children understand the intentional actions, including communicative actions, of other persons.

The process is never totally straightforward. The philosopher Wittgenstein (1953) noted that even an ostensive definition – the seemingly simplest case of language acquisition in which one person “shows” another what a word means – is problematic because it assumes that both teacher and learner know what “showing” is and precisely how it serves to pick out individual referents in some language-independent way. The point was crystalized by Quine (1960) in his parable of a native who utters the expression “Gavagai!” and “shows” a foreigner the intended referent by pointing out a salient event as it unfolds. Given the stipulation that native and foreigner have no way to establish a common view of the event nonlinguistically, however, there is basically no way that the foreigner can know whether the native’s novel expression is being used to refer to the event, to some participant in the event, to some part of the participant’s body, to the color of the participant’s hair, or to any of an infinite number of aspects of the situation. This is the basic problem of referential indeterminacy.

In the modern study of language acquisition there are essentially two approaches to this fundamental problem of language acquisition: the constraints approach and the social-pragmatic approach. In this chapter my intention is to provide evidence for the social-pragmatic approach, and in particular for the view that to acquire linguistic conventions in the situations in which they encounter them, young children must have more powerful skills of social cognition than is generally recognized. In all cases the key

social cognitive skill is children's ability to perceive the intentions of the adult as she acts and speaks to the child. In making my case I first outline the two predominant views of early word learning; I then summarize seven experimental studies in which 18- to 24-month-old children try to determine adults' referential intentions in complex interactive situations; and I conclude by arguing more generally that the understanding of intentions – specifically, the understanding that other persons have intentions towards my intentional states – is the very foundation on which language acquisition is built.

1 Two views of word learning

One approach to the problem of referential indeterminacy in the study of lexical acquisition is the so-called “constraints” approach (e.g. Markman 1989, 1992; Gleitman 1990). In this view a learner, who knows already what words are, attempts to acquire a new word by: (1) creating a list of possible hypotheses about how the new word “maps” onto the real world, and (2) eliminating incorrect hypotheses in a semi-scientific manner. In all cases the learner's goal is to perform the task correctly and so to acquire accurately the mapping between word and world. The problem is that, as the philosophers have pointed out, there are simply too many possible hypotheses to be tested in a given case. Therefore, the child must be given a headstart on the process. This headstart takes the form of word learning “constraints” that eliminate certain hypotheses before they are seriously entertained.

One possible constraint on the process is that young children may be biased to pay attention to only some things in the environment, and they may know prior to any language learning that these will be the first referents of adult words addressed to them. For example, in Markman's (1992) account, young children attend to whole objects more than to their parts, properties, or activities, and so they assume initially that novel words refer to whole objects. In this theory, children also know prior to language acquisition that two different words do not map onto the same real world object (mutual exclusivity); therefore, if an adult uses a new word in the presence of an object whose name they already know, children assume that something other than the name of that whole object is the word's referent. The theory does not specify how children determine precisely which other aspect of a situation is being referred to by a word that is not the name of a whole object (although see Golinkoff, Hirsh-Pasek, Mervis, *et al.* 1995, for some suggestions). Other constraints theorists believe that for some classes of linguistic item, especially verbs and other predicative terms, the key to lexical acquisition is information provided by children's *a priori* knowledge of syntax (syntactic bootstrapping; Gleitman 1990). In this view, the syntactic structures

surrounding words, verbs in particular, provide additional constraints on their possible mappings to the world of events and states of affairs. Although no one has to my knowledge proposed a conjoining of the two theories, it is perhaps possible that children use both word learning biases and syntactic information to help constrain their hypotheses about the meaning of novel linguistic items and expressions.

By all accounts – including those of Markman and Gleitman – these word learning constraints cannot solve the problem of referential indeterminacy by themselves. They only work if used in conjunction with social-pragmatic information about another person's specific referential intentions in specific contexts. That is, even if the child knows that an adult is using a new piece of language to refer to a whole object or a certain type of event, there is still the problem that in many word learning situations there are multiple referents in the immediate context that fit within these specifications. One way to supplement the constraints account is thus to add a few very specific pragmatic cues that, in combination with the various word learning constraints, would be sufficient to enable the child to determine the referential significance of novel words. One obvious candidate is something like eye gaze direction. For example, in this account a child might know *a priori* that an adult is referring to a whole object, but would use gaze direction to specify which of several possible candidates is the intended referent (as in the studies of Baldwin 1991, 1993a). The problem with this supplemental account is simply that gaze direction is itself insufficient in many cases of word learning, as adults quite often talk to children about referents that are not visually accessible at all – for example, in referring to an absent object or in requesting actions that the child may or may not actually perform. And children do learn words in these contexts as several empirical studies demonstrate (Tomasello 1992a; Tomasello & Kruger 1992; and all of the studies to be reported in this chapter).

The social-pragmatic approach to the problem of referential indeterminacy takes a very different perspective on lexical acquisition. It begins by rejecting truth-conditional semantics in the form of the mapping metaphor (the child maps word onto world), adopting instead an experientialist and conceptualist view of language in which linguistic symbols are used by human beings to invite others to experience situations in particular ways. Thus, attempting to map word to world will not help in situations in which the very same piece of real estate, for example, may be called: *the shore* (by a sailor), *the coast* (by a hiker), *the ground* (by a skydiver), and *the beach* (by a sunbather) (Fillmore 1982; Langacker 1986; Lakoff 1987). In the social-pragmatic view, each of the world's natural languages has its own set of communicative conventions, in the form of linguistic symbols created over thousands of years of human history, by means of which its speakers attempt to influence the interest and attention of other members of their

speech communities (Talmy 1996). There are also universals in the way symbols are created, learned, and used across languages, of course, reflecting universals both in the way human beings experience the world and in the ways they interact and communicate with one another socially. There is nothing in human languages other than these symbols – which include grammatical symbols and categories of symbols – and the experiences they symbolize (Langacker 1987, 1991).

In the social-pragmatic approach to language acquisition in general, and to referential indeterminacy in particular, the focus is on two aspects of the process: (1) the structured social world into which the child is born – full of scripts, routines, social games, and other patterned cultural interactions; and (2) the child's capacities for tuning into and participating in that structured social world (Tomasello 1992b). Language is one means by which adults exhort children to attend to certain aspects of a shared social situation. In attempting to comply with these exhortations – that is, in attempting to comprehend adult use of these symbols and so to have the requested experiences – children use all kinds of interpretive strategies based on the pragmatic assumption that adult linguistic symbols are somehow *relevant* to the ongoing social interaction (Bruner 1983; Sperber & Wilson 1986; Bloom 1993). In the social-pragmatic view, young children are not engaged in a reflective cognitive task in which they are attempting to make correct mappings of word to world based on adult input, but rather they are engaged in social interactions in which they are attempting to understand and interpret adult communicative intentions – so as to make sense of the current situation (Nelson 1985). Having complied with adult instructions to experience a situation in a particular way in a given instance, children may then learn to produce the appropriate symbols for themselves when they wish for others to experience a situation in that same way – thus entering into the world of bidirectionally (intersubjectively) understood linguistic symbols (Tomasello 1996).

In the social-pragmatic view, then, children acquire linguistic symbols as a kind of by-product of social interaction with adults, in much the same way they learn many other cultural conventions (Tomasello, Kruger, & Ratner 1993). The acquisition of linguistic symbols does not need external linguistic constraints in this theory because children are always participating in and experiencing particular social contexts, and it is these social contexts that serve to “constrain” the interpretive possibilities. The child who knows that his mother wishes him to eat his peas (she is holding them up to his mouth and gesturing) assumes that her utterance is relevant to that intention, and this is what guides his interpretations of any novel language in the utterance. All of the philosophically possible hypotheses that Quine and others may create are simply not a part of the child's experience in this particular social context – assuming of course that by the time language acquisition begins

young children do indeed have a reasonably adult-like understanding of at least some aspects of the social activities in which they participate. And of course a prerequisite for word learning is the child's ability to conceptualize referents in a manner similar to adults; it is just that in the social-pragmatic view these conceptualizations are not tied to language *a priori*, as they are in the constraints view, but rather the child must learn the connections from participating in meaningful communicative interactions.

Some early studies in the social-pragmatic perspective focused on the role of the adult in the acquisition process. Thus, it was found that when adults name new objects for young children by following into their already-established focus of attention, as opposed to using the new language to direct their attention to something new, word learning is made easier (Tomasello & Farrar 1986; Dunham, Dunham, & Curwin 1993). But this does not mean that the child is a passive participant in the word learning process. Subsequent studies have shown that when there is a discrepancy between the focus of attention of adult and child – when young children hear a novel word in situations in which their focus of visual attention differs from that of an adult – learners nevertheless are able to do the extra work to determine the adult's referential intentions, almost never assuming that the new word is being used for whatever is their current focus of attention irrespective of what the adult is attempting to do (Baldwin 1991, 1993a). In all cases of word learning, children make active attempts to understand adult referential intentions; it is just that some situations make it easier or harder for them to do so.

My own biases, as alluded to above, are towards the social-pragmatic view. It seems to me that children do not learn words by employing mechanical mapping procedures or testing hypotheses, but rather that they learn them in the same basic way they learn other cultural skills and conventions: in the flow of naturally occurring social interaction in which both they and their interlocutors have various pragmatic goals towards the world and towards one another. Adults may do things that facilitate the learning process, but the child must always do some social cognitive work to determine the adult's referential intentions. Children learn words as an integral part of their social interactions with other persons, an important part of which are their attempts to understand what adults are trying to get them to do and their own attempts to get adults to do things (Bruner 1983; Nelson 1985; Tomasello 1992b).

2 Learning words “in the flow” of social interaction

In my view, one of the main reasons that researchers have been drawn to at least some aspects of the constraints view is that the prototypical case of

word learning is assumed to be the learning of an object label in an ostensive context: an adult intends that a child learn a word and so “shows” her its referent (e.g. by holding up or pointing) in temporal contiguity with the utterance. In this case, the pragmatic understanding involved – that the adult intends to indicate the object held up or pointed to – is so basic and seemingly simple that it is often overlooked completely. However, when we turn our attention to the acquisition of words with more relational meanings – for example, verbs and prepositions in English – this picture of the basic word learning situation changes. Western middle-class adults use verbs to children most often to regulate or anticipate their behavior, not to name actions for them (Tomasello & Kruger 1992). In such cases, the pragmatic cues that might indicate the adult’s intended referent are less straightforward than in the ostensive context, and indeed they change in fundamental ways from situation to situation: the adult requests for the child to eat her peas by directing the spoon at her face, but requests that the child give her something by holding out her hand. That is to say, there is no standardized “original naming game” for actions/verbs, as there is for some children learning object labels (Tomasello 1995a). In other cultures, the prototypical word learning situation is even more dramatically divergent from the classic ostensive context, as adults rarely stop what they are doing to name things for children at all, and in some cases speak directly to them only infrequently (Brown and de León, chs. 17 and 18, this volume).

It would thus seem from casual observation that young children are able to learn words in a wide variety of social-interactive situations that differ significantly from the ostensive context. In the studies that follow I attempt to demonstrate experimentally that this is indeed the case for children in their second year of life – that is, for children who have just caught on to the process of word learning and have begun to build their initial vocabularies. Moreover, by establishing something of the wide variety of situations in which these nascent language learners can acquire new words, I attempt to undermine any attempt to characterize the process as wedded to any specific pragmatic cues such as gaze direction. Instead, I argue that learning new words is dependent on young children’s ability to perceive and comprehend adult intentions in a very flexible manner using many different types of social-pragmatic information. A result of this line of argumentation is that the process of word learning comes to be seen as, in essence, just another manifestation of processes of cultural learning in general, and, like other manifestations of cultural learning, it depends on children’s basic skills of social cognition (Tomasello *et al.* 1993).

The experimental studies my collaborators and I have recently performed share a number of features. The basic idea in all cases is to set up situations in which adults talk to children as they engage in various games, with novel

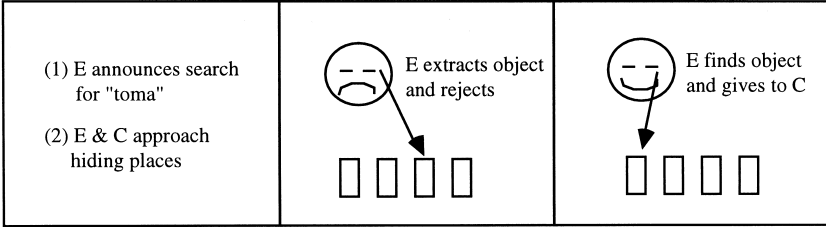


Fig. 5.1 The With Search condition in Tomasello & Barton (1994: Study 4).

words being introduced as naturally as possible into the ongoing flow of the game. In all cases there are multiple potential referents available; that is, there are multiple novel referents for which the child has no existing means of linguistic expression (this is checked before each study begins), and the novel word is introduced in a single type of linguistic context. Various pragmatic cues to the adult’s intended referent are provided in different studies to see if children are indeed sensitive to them. The studies are designed so that none of the well-known word learning constraints that various investigators have proposed (e.g. whole object, mutual exclusivity, syntactic bootstrapping) will be helpful to the child in distinguishing among possible referents. The studies are also designed so that eye gaze direction is never diagnostic of the adult’s referential intentions.

2.1 *Determining which object the adult intends to find*

In our first study of this type, Tomasello & Barton (1994, Study 4) had a female experimenter say to 24-month-old children, “Let’s go find the toma,” while looking directly into the child’s eyes. The two of them then approached a row of five buckets. Each bucket contained a novel object, none of whose names the child knew beforehand, with the target object designated randomly across children. There were two experimental conditions. In the Without Search condition the adult went immediately to a bucket and excitedly found the target object and handed it to the child. In the With Search condition the adult went to the buckets, first extracted and rejected (by scowling at and replacing) two objects, and only then excitedly found and handed the child the target object (see figure 5.1). In both conditions the finding event was followed by the excited extraction of each of the other objects as the adult said “Let’s see what’s in this one.” After several rounds of this procedure, comprehension and elicited production testing were conducted. The comprehension test consisted of the adult motioning in the general direction of all five objects (laid out in random order by a research

assistant) and asking the child to bring her the toma. The elicited production task consisted of the adult holding up the target object and asking the child for its name (spontaneous productions at other times during the experiment were also recorded).

The outcome was that children learned the new word equally well in the two experimental conditions, both in comprehension and in production. Our interpretation of this result was as follows. Children could not have used a word learning cue such as “the object at which the adult is looking while saying the new word” because the adult was always looking into the child’s eyes at this time. Nor could they have used some simple extension of this cue such as “the first object the adult looks at after uttering the new word” (or “first object the child sees”), because then they would have performed better in the Without Search condition in which the target object was the first one extracted from the buckets; moreover, if they had used this cue in the With Search condition, they would have thought that the first object extracted and rejected was the toma (since the same object was always the first rejected object in all models for a given child) – which they did not, as they did not in a similar study by Baldwin (1993b). Because children performed equally well in the two conditions, our assumptions were that: (1) the children understood from the beginning that the adult’s intention was to find a particular object called a toma; and (2) they tracked the adult’s behavior and emotional expression until she seemed to fulfill that intention by expressing excitement and terminating the search (i.e. with the first object extracted in the Without Search condition and the third object extracted in the With Search condition). We do not know precisely which aspect of the adult’s behavior was critical for the children – her excited demeanor or the termination of the search, for example – but in either case the adult’s behavior was a meaningful cue precisely because it was an indication that she had fulfilled her intention to find the toma.

Akhtar & Tomasello (1996, Study 1) had 24-month-old children participate in a variation of this game. There were two main differences. The first was that one of the buckets in the row of five was actually a very distinctive toy barn (again there was a novel, nameless object hidden in each container). The second was that there were some initial rounds of a finding game (in which the adult used no new words), so that the child would come to know which object was in the toy barn (randomized across children). The procedure was thus as follows. The adult and child went through the buckets and barn several times extracting the objects, with the adult saying things like “Let’s see what’s in here.” After these initial extractions, the language models began, with the locations of all objects remaining constant for a given child and the object in the barn always serving as the target object. In the language modeling rounds, the experimenter announced her

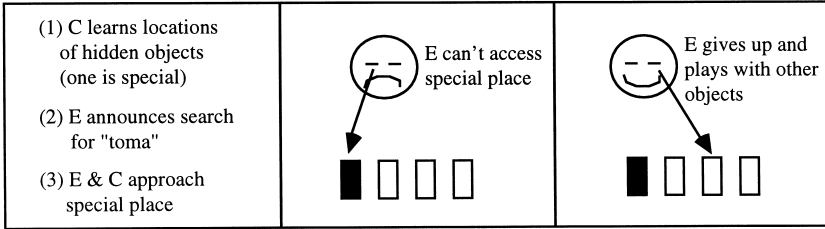


Fig. 5.2 The Absent Referent condition in Akhtar & Tomasello (1996: Study 1).

intention to find a specific object: “Now let’s find the toma!” There were two experimental conditions. In the Referent condition, the experimenter proceeded directly to the barn and extracted the target object (then followed by extraction of the other objects). In the Absent Referent condition, she proceeded to the barn and attempted to open it, but after being unsuccessful (and looking disappointed) had to inform the child “It’s locked. I can’t open it” (followed by extraction of the other objects; see figure 5.2). Thus, children in the Absent Referent condition never saw the target object after hearing the novel word. Language modeling was followed in both conditions by comprehension testing and elicited production.

Once again the outcome was that children learned the new word equally well in the two experimental conditions. (As an additional control to make sure that children were actually learning the word additional subjects also participated in versions of the experimental conditions without the target language – the adult simply said of each container “Let’s see what’s in here” – and children performed at chance in the comprehension test.) As in the Tomasello & Barton study, this pattern of results indicated to us that (1) the children understood that the adult’s intention was to find the toma, and (2) they then tracked her behavior and emotional expression in an attempt to determine when she had succeeded in doing so. In the Absent Referent condition, however, the adult did not succeed and thus never showed excitement in finding the target object – indeed she expressed her disappointment explicitly. Therefore, not only could children not have been using adult gaze direction, or “next object the adult sees,” in this condition, they also could not have been using the adult’s excitement as a cue; when the barn was locked the adult showed nothing but disappointment. Indeed, if they had used adult excitement as a cue, children would have inferred that the toma was the object extracted excitedly from the adjacent hiding place right after the failure, and almost no children did this. In this study children understood the adult’s finding intention and used as a word learning cue her dis-

appointment at her failure, remembering or imagining in the process what it was she had intended to find.

Taken together these two studies demonstrate that children have multiple ways of understanding adult intentions in finding games. Children come to the experiment knowing about finding as an intentional activity. They soon come to understand that the adult intends to find this thing called a toma. They then monitor the adult's behavior to see when she has fulfilled her intention. Across the two studies they must do this in different ways: in one study understanding that the object the adult is excited about is the toma, while in the other study understanding that the object never seen – the one the adult is disappointed in not finding – is the toma. The implication is thus that in both of these studies children are employing a very flexible social understanding of adults' actional and communicative intentions and how they play themselves out in different circumstances. Recently Tomasello, Strosberg, & Akhtar (1995) have extended these two findings to children eighteen months of age, demonstrating that these intentional understandings are not some late-developing word learning strategy, but rather are an integral aspect of children's word learning skills from very early in language development.

2.2 *Determining which action the adult intends to perform*

From the beginning of our thinking about the role of intentional understanding in word learning, the acquisition of verbs was seen as especially important. This is because many verbs (especially change-of-state verbs such as *give*, *put*, *make*, etc.) are only heard by young children as the adult requests an action of the child, anticipates the child's or her own impending action, or comments on a completed action – basically never using them in ostensive naming contexts (Tomasello & Kruger 1992). Verbs are thus the prototype of words that are experienced by children in the ongoing flow of social interaction with others.

In our first experimental study along these lines, Tomasello & Barton (1994, Study 3) exposed 24-month-old children to two novel verbs, one in each of two experimental conditions, both involving an impending action on an apparatus affording two actions. In the Target First condition an adult experimenter used a novel verb to announce her intention to perform a novel action on the apparatus, saying, for example, "I'm going to plunk Big Bird!" She then performed the target action intentionally, saying "There!", followed immediately by another action on that same apparatus performed "accidentally," in an awkward fashion, saying "Whoops!" In the Target Last condition, the accidental action was performed first and the intentional action was performed second (see figure 5.3). (Because there

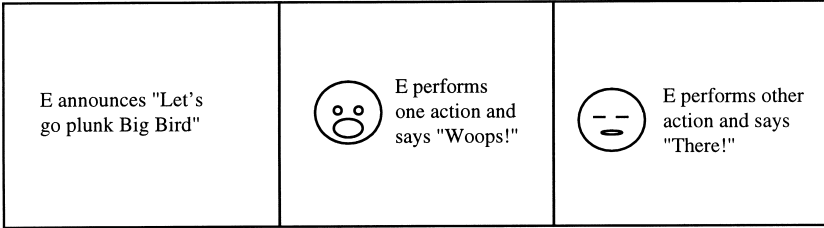


Fig. 5.3 The Target Last condition in Tomasello & Barton (1994: Study 3).

were two apparatuses, each with two possible actions, which actions were intentional and which were accidental – as well as the order of intentional and accidental actions – was totally counterbalanced across children.) After several models of this type, the child was given a new character and asked “Can you go plunk Mickey Mouse?”, with both apparatuses present. There were also attempts to get the children to produce the new word as the adult performed the action.

Results demonstrated that the children quite readily associated the new action word with the adult’s intentional action, regardless of whether it occurred immediately after the novel word or later, after an intervening accidental action. As in the previous studies, the exact cues the children might have been using to distinguish intentional and accidental actions were not identified precisely; the adult both acted differently and used a different vocal marker for the two types of action. But again, as in the previous studies, whatever cues they might have used, the children in this study apparently understood that the adult intended to plunk Big Bird, and they then tracked her behavior and emotional expressions to discover precisely what this plunking action might be.

Related to this finding is a study in which Akhtar & Tomasello (1996, Study 2) had an experimenter expose 24-month-old children to four novel actions using no new words. Each of these actions was uniquely associated with one and only one toy character and one and only one prop (e.g. catapulting was associated only with Ernie and only with a device for catapulting). Then came several language modeling rounds using a novel verb (the target) associated with one of the four actions (counterbalanced across children). There were two experimental conditions. For subjects in the Referent condition, the experimenter set out the target action prop and announced her intention to perform an action – for example “Let’s pud Ernie!” She then proceeded to perform the target action (then followed by the other actions with their appropriate props and characters). For subjects in the Absent Referent condition, the experimenter also set out the target action prop and announced the same intention but, after searching, told the

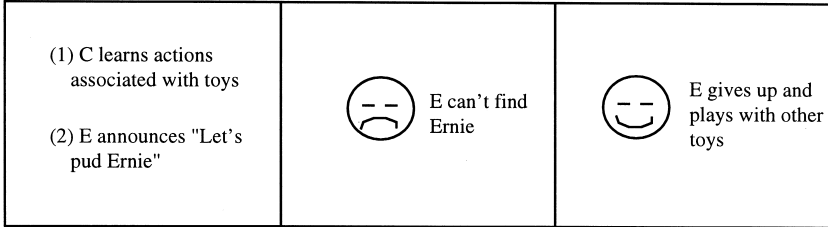


Fig. 5.4 Absent Referent condition in Akhtar & Tomasello (1996: Study 2).

child she was unable to find Ernie and so she could not pud him (so she then went on to bring out the other props and characters and perform their respective actions; see figure 5.4). Children in this condition thus never saw the referent action after hearing the novel verb. Language modeling was followed in both conditions by comprehension testing and elicited production.

As in the very similar study with objects (see above), the outcome was that children learned the new word equally well in the two experimental conditions. (As an additional control to make sure that children were actually learning the word, additional subjects also participated in versions of the experimental conditions without the target language – the adult simply continued to say “Let’s do this” – and they performed at chance in the comprehension test.) As in the similar study with objects, this pattern of results indicated to us that (1) the children understood that the adult’s intention was to pud Ernie, and (2) they then tracked her behavior in an attempt to determine when she had done so. In the key condition, however, the adult never actually performed the action. Indeed, in the Absent Referent condition the next action the adult performed (after the disappointment of not being able to perform the target action) was one of the non-target actions. Thus, as in the similar study with objects, children in this study were able to understand in this highly scripted situation what the adult intended to do from her disappointment, without ever seeing the referent action paired with the novel verb.

These two studies with novel verbs thus demonstrate rather directly the role of children’s understanding of intentional action in early word learning. Children know that adults use words to announce their intended actions, not their accidental actions (which would, of course, make no sense), and they have some ability in highly scripted situations to anticipate the adult’s intended action even when it never actually happens and the adult shows disappointment – again evidencing a very flexible social understanding. We have not attempted these studies with younger children. However, two recent studies in which 16- and 18-month-old infants were

asked to imitate novel actions on objects suggest that even at this younger age the ability to understand intentional action is present. In perfect parallel to the two verb-learning studies just reported: (1) Carpenter, Akhtar, & Tomasello (1998) found that 16-month-old infants were much more likely to imitate an adult's intentional than her accidental actions (cued mainly with the words "Whoops!" and "There!" said with their appropriate intonations); and (2) Meltzoff (1995) found that 18-month-old infants were able to "imitate" full actions that the adult only initiated but did not complete (and thus the intention was never actually fulfilled). It is thus possible that children this young could learn novel verbs in our two experimental paradigms as well.

2.3 *Determining what is new for the adult*

In each of the studies just reported there was some set of pragmatic cues indicating which object or action the adult intended to indicate with her novel word. In a fifth study attempting to demonstrate the role of children's intentional understanding in the word learning process, Akhtar, Carpenter, & Tomasello (1996) set up a situation in which no special cues were associated with any one referent. In this study, 24-month-old children first played with three novel, nameless objects with two experimenters and a parent. The adults drew attention to each of the objects, handing them to the child and to one another excitedly, commenting on their characteristics, and in general playing with the objects enthusiastically. No language models were given during this initial play period. One experimenter and the parent then left the room while the child and the other experimenter played with a fourth object (the target) for an amount of time equal to that for which they had played with the other objects. The experimenter then set all four of the objects in a row, in random order. The other adults then returned. What they then did constituted the two experimental conditions. In the Language condition the adults looked at the group of objects (none singled out by gaze) and said: "Look, I see a gazzer! A gazzer!" (see figure 5.5). In the No Language condition the adults behaved the same but said: "Look, I see a toy! A toy!" (There was also another control study in which the novel object was not the last one the child played with.) In both conditions, the language model was later followed by a comprehension test in which the experimenter who had left the room asked the child to bring her the gazzer, and there was an elicited production question as well.

The main result of this study was that children learned the word for the target object in the Language condition, but they behaved randomly in the No Language condition. This result means that the children: (1) knew which object was novel for the adults who had left the room; and (2) knew

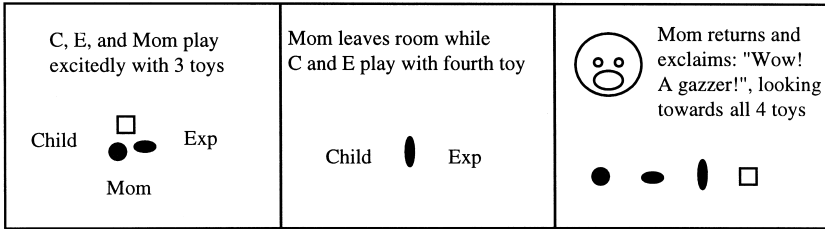


Fig. 5.5 The Language condition in Akhtar, Carpenter, & Tomasello (1996).

that adults only get excited about, and thus only use new language to talk about, things that are new to the discourse context. (Another way of expressing this second point is to say that the children knew that the adult had played with the other toys repeatedly previously and so it did not make sense for her to be excited about seeing them now, thus eliminating them as possible referents – analogous to the process of elimination in so-called “fast mapping” studies.) The important point is that the children in this study were able to single out the target object not on the basis of its being treated differently by the adults giving the language model; the pragmatic “cue” in this case was something much more distributed in terms of the child’s understanding of the adults’ experience during the entire experimental situation.

The power of this kind of understanding is underscored by another of our recent studies. Tomasello & Akhtar (1995, Study 1) were interested in whether children in this same general age range (24–26 months) could use their understanding of novelty from the adult point of view to help them determine to which ontological category a novel word might belong. The basic idea was that the language model would be given identically for all children – that is, using exactly the same language in exactly the same referential situation. What would differ across experimental conditions would be the experiences children had leading into that language model. There were two experimental conditions. In both conditions, children heard a novel word modeled as a one-word utterance (“Modi!”) just as a nameless target object was engaging in a nameless target action. What preceded this model defined the experimental conditions. In the Action Novel condition, children initially performed several actions on the target object, so that when the language model was presented the target action was the novel element in the discourse context. (The “feel” was supposed to be something like: first we do this with it, then we do that with it, now we “modi”.) In the Object Novel condition, children performed the target action on several objects so that when the language model was presented the target object

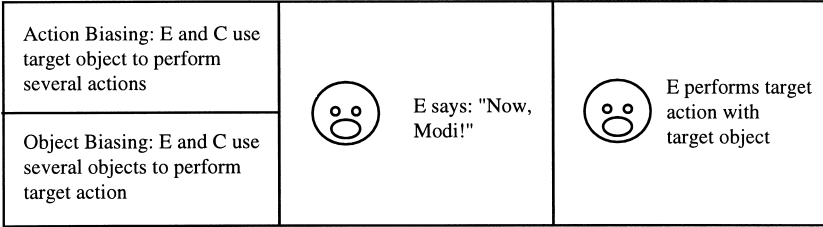


Fig. 5.6 Both experimental conditions in Tomasello & Akhtar (1995: Study 1).

was the novel element in the discourse context. (The “feel” was supposed to be something like: first we do it with this, then we do it with that, now we do it with “modi”.) Figure 5.6 depicts both experimental situations.

There was an elicited production question and a comprehension test. The comprehension test was preceded by a pre-test (which actually took place at the very beginning of the experiment) in which children were given a cup, a spoon, and a ball. They were then asked to “Show me ball,” “Show me bounce,” and so forth, as a way of letting them know that with this one sentence frame (“Show me ____”) we might be asking for either an object or an action. Most children showed clearly from the beginning a tendency to hold up or give objects when they were requested, and to perform actions when they were requested. (Any children who did not were trained to do so, or, in a few cases, dropped from the study.) The target comprehension test was thus “Show me modi” addressed to the child in the presence of multiple objects (all of which had been played with equally) and multiple possible actions (all of which had been performed equally). Results indicated that the children associated the word *modi* with the element that was new to the discourse context at the time of the model: children in the Action Novel condition associated *modi* with the target action, whereas children in the Object Novel condition associated *modi* with the target object. This finding is especially significant because it indicates for the first time that children’s intentional understanding can even take them across ontological categories. That is, in all of the previous studies reported, children have understood from the outset that the adult was referring to either an object or an action and their task was to determine *which one*; in this study the child’s task was to determine *what kind* of entity the adult intended to indicate with her novel language.

These two studies thus show that young children can use something like novelty to the discourse situation to determine adult referential intentions. In the first study children had to determine which one of a number of objects was new, and they had to do this totally from the adult’s point of view. In the

E and C play with target object engaged in target action (no language)	Action Biasing: E prepares apparatus for C, alternates gaze between apparatus and C	Action Biasing: E says: "Your turn, Jason. Wigit!"
	Object Biasing: E ignores apparatus, alternates gaze between object and C	Object Biasing: E says: "Wigit, Jason! Your turn."

Fig. 5.7 Both experimental conditions in Tomasello & Akhtar (1995: Study 1).

second study novelty was equal for both child and adult, but children had to use this novelty to determine what kind of entity, object or action, the adult intended to indicate. The ability to determine what is new in a situation for a potential communicative partner is of course not only an important skill in early word learning, it is also a crucially important skill in all aspects of language development as children learn to comprehend and produce utterances that are pragmatically appropriate to specific communicative contexts (e.g. for all kinds of topic-maintenance operations in discourse).

2.4 *Determining what the adult intends for me to do*

The final study in this series is Tomasello & Akhtar (1995, Study 2). The basic design is similar to that of the study just described: children saw a novel word modeled in a situation in which the adult might potentially be referring to either an object or an action. In this case, however, the adult behaved in a number of ways that differed between the two experimental conditions. To begin, all children saw an adult perform a novel and nameless action (target action) with a novel and nameless object (target object) on a special apparatus, and then they had the opportunity to perform the action with the object themselves – with no new words used at this time. Children in the Action Highlighted condition then watched as the adult prepared the apparatus so that the child could perform it again (by orienting the apparatus correctly for him). She then held out the object to the child and said “Your turn, Jason. Wigit!”, while alternating her gaze between the child and the apparatus – as if requesting that the child perform the action. The experimenter behaved differently for children in the Object Highlighted condition. In this condition the experimenter did not prepare the apparatus for the child, and simply held out the object to the child and said “Wigit, Jason! Your turn,” while alternating her gaze between the child and the object (never looking at the apparatus at all). Figure 5.7 depicts both of these conditions.

Elicited production and comprehension tests were then given, with the “Show me wigit” test, prepared by a pre-test, used to test comprehension as reported above (i.e. as in Tomasello & Akhtar, 1995, Study 1). The outcome was that children in the Action Highlighted condition learned the new word for the target action, whereas children in the Object Highlighted condition learned the new word for the target object. Once again a key aspect of this study was the fact that children did not know ahead of time whether we would be talking about a novel object or a novel action; it was only during the modeling itself that they learned the experimenter’s referential intentions with respect to ontological category. The experimenter in this study did several things that differed in the two conditions: either prepared or did not prepare the apparatus, alternated gaze between the child and either the apparatus or the object, and said either “Your turn, Jason. Wigit!” or “Wigit, Jason! Your turn.” Again in this case, therefore, we do not know the precise cue or cues used by children to determine the adult’s referential intentions, but we do know that they used some aspects of the presented cue complex to distinguish situations in which the adult intended for them to perform an action from situations in which the adult intended to name an object for them.

It would also be interesting to perform this study, or something like it, with younger children. We have no experimental evidence that they could learn new words in this situation, but in Tomasello (1992a) I reported my daughter’s acquisition of several verbs that seemingly were used by adults only in requestive situations (she sometimes complied and sometimes did not). A reasonable hypothesis is thus that even younger children could also learn novel verbs (at least in comprehension) from utterances in which adults request behaviors of them even if the actions are never actually performed – so long as the appropriate intention cues are present.

2.5 Summary

The fact that children can learn words in all of these different interactive situations – none of which consists of the adult stopping what she is doing to name something for the child – is a very important fact for theories of word learning (see table 5.1 for a summary). There are at least two possible explanations for this fact. On the one hand, it is possible that young children learn to deal with each of these learning situations separately – that is, they learn in each situation separately how the adult’s new word “maps” onto one of the potential referents available: for example, through some kind of hypothesis-testing procedure with general *a priori* constraints, individual cue learning, and subsequent feedback. On the other hand, it is also possible that word learning is not a hypothesis-testing procedure needing to be

Table 5.1. *Some of the different sources of information by means of which children determine adult referential intentions in the second year of life – as evidenced by studies of word learning using experimental techniques*

Information from adult	Study
1. Directedness to an Object (including eyes, body, and voice)	Baldwin (1991, 1993a)
2. Expression of Intent and its Satisfaction (both objects and actions)	Tomasello & Barton (1994)
3. Expression of Intent + Event Knowledge (both objects and actions)	Akhtar & Tomasello (1996)
4. Expression of Surprise (new element for adult)	Akhtar <i>et al.</i> (1996)
5. Expression of Surprise (new element for adult across object–action boundary)	Tomasello & Akhtar (1995)
6. Expression of Intent Towards Child's Actions (across object–action boundary)	Tomasello & Akhtar (1995)

constrained at all, but rather it is a process of skill learning that builds upon a deep and pervasive understanding of other persons and their intentional actions (i.e., social cognition in general) that is available to children by the time language acquisition begins. This is the hypothesis to which we now turn.

3 Learning the first word

All of the studies just reported concern children in the 18- to 24-month age range. This is the period during which many children have a “vocabulary spurt” and thus would seem to be the period during which they have become skillful at learning new words – to the tune of several new words per day for many children. It is thus possible that what these studies represent is a set of acquired strategies for word learning that only become operative after a certain developmental period. In this scenario children’s acquisition of their first words in the first half of the second year of life might not depend on an understanding of the intentional actions of others in the same way as does their word learning later in the second year. The other possibility is that, perhaps with some minor twists, the basic process of word learning is the same throughout early development. Indeed, in this view, the creation of linguistic conventions/symbols in the first place can only occur in organisms who understand their communicative partners as intentional agents. Children’s learning of their first words is thus of a piece

with their other social cognitive and cultural learning skills at this same developmental period early in the second year. This does not mean that word learning does not involve some additional complexities, only that it is not a completely modularized process in early development.

The story goes something like this. Human infants are social creatures from birth and so they clearly understand many things about other persons early in development. But they do not seem to understand other persons in terms of their intentional relations to the world. Five-month-old infants, for instance, do not do such things as following the gaze of others to outside entities or imitating their behavior on outside objects, preferring instead to focus on their own face-to-face interactions with others. All of this changes, however, at around the first birthday. Infants at this age begin for the first time to do such things as looking towards objects at which adults are looking (gaze following), monitoring adult emotional reactions to novel objects (social referencing), and acting on objects in the way adults act on them (imitative learning). All of these skills emerge in rough developmental synchrony because all of them reflect the infant's emerging ability to understand other persons as intentional agents whose attention, emotion, and behavior to outside objects may be actively followed into and shared (Tomasello 1995b). Experimental support for this view has recently been provided by the studies of Gergeley, Nádasdy, Csibra, & Biró (1995).

This social cognitive revolution at the infant's first birthday sets the stage for the second year of life in which it begins to learn imitatively the use of all kinds of tools and artifacts, with linguistic symbols being one special case. In all of these cases the imitative learning involved is not just a mimicking of adult body movements, and not just a reproducing of interesting environmental effects by whatever means imaginable, but rather it is an actual reproduction of the adult's intentional relations to the world. In these interactions the infant perceives the adult's overt actions as composed of both a goal and a means for attaining that goal, and then actively chooses the adult's means of goal attainment in contrast to others it might have chosen. For example, in a study by Meltzoff (1988) 14-month-old children observed an adult bend at the waist and touch her head to a panel, thus turning on a light. Most of them imitatively learned this somewhat awkward behavior, even though it would have been easier and more natural for them simply to push the panel with their hands. The supposition is that the infants understood that (1) the adult had the goal of illuminating the light, (2) chose one means for doing so, from among other possible means, and (3) if they had the same goal they could choose the same means. Imitative learning of this type thus relies fundamentally on infants' tendency to identify with adults, and on their ability to distinguish in the actions of others the underlying goal or intention and the different means that might be used to achieve it.

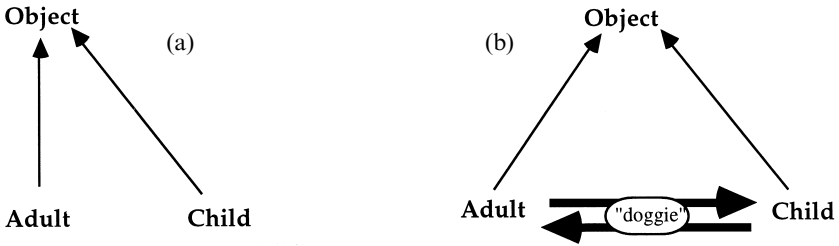


Fig. 5.8 (a) Adult acts on object and child imitatively learns that action. (b) Adult uses word to express intention (thick line) towards child's attention (thin line to object), and child imitatively learns that word with role reversal, thus creating a bidirectional linguistic convention/symbol.

This interpretation is supported by the recent finding of Meltzoff (1995) that 18-month-old infants also imitatively learn actions on objects that adults intend to perform, even if the adult is unsuccessful in actually performing them. It is likewise supported by Carpenter *et al.* (1998), who document that 16-month-old infants imitate adults' intentional actions on objects (indicated by "There!") and ignore their accidental actions (indicated by "Whoops!") – in perfect parallel to the study of Tomasello & Barton (1994, Study 3 – as reported above).

The acquisition of linguistic symbols begins during this same developmental period quite simply because comprehending and producing language rely on the same basic understanding of persons as intentional agents as all of the other social cognitive and cultural learning skills that emerge at this same age. Nevertheless, linguistic symbols do have a special property. Unlike actions on objects, words are communicative actions directed to other persons; that is, words are social conventions and so involve a special application of social cognitive and cultural learning skills. This can be made clear by comparing the process by which infants imitatively learn an action on an object and the process by which they imitatively learn a linguistic convention.

Let us suppose that an infant observes an adult performing an intentional action on an object, and then imitatively learns that action herself. The process may be depicted as in figure 5.8a. Now suppose that an adult addresses a novel expression to a child, perhaps while pointing to an interesting event – for example, "A doggie!" The key is that the adult's intentional action in this case is not directed to the object, but rather to the child or, more specifically, to the child's intentions or attention. The child's comprehension of the adult's communicative act, therefore, requires an understanding that the adult is making this unfamiliar noise with the goal of getting me to do something intentional (i.e. focus my attention on some

aspect of the ongoing event; see Gibson & Rader 1979, for the argument that attention is intentional perception). Once the child understands this communicative intention, a linguistic convention is created when she then acquires appropriate use of that communicative act herself. This process is similar to other acts of cultural learning in that the child understands and reproduces the adult's intentional act. It is different, however, in that the imitative learning of communicative acts involves a role reversal. If the child understands the adult's act as "She is making that sound in order to get me to attend to the doggie," then imitatively learning this act means that the child must reverse the roles: if *I* wish *others* to focus on the doggie *I* must use the sound "doggie" towards *them* (see figure 5.8b). As the child identifies with the adult and adopts her linguistic expression, of course, she retains the understanding that the adult also comprehends and uses it. The child's use of the expression thus creates a communicative convention, or symbol, whose essence is its bidirectionality or intersubjectivity (Saussure 1916), which constitutes the quality of being socially "shared" (Akhtar & Tomasello, in press).

It is important to note that not all communicative behaviors are bidirectional and shared in this same way. Thus, prelinguistic infants use a number of communicative gestures that are ritualized from noncommunicative behaviors; for example the "hands-up" gesture as a request to be picked up may be a ritualization of the infant trying to pull its way up to the parent's arms. There is no evidence that prelinguistic infants learn this gesture by imitative learning, or even that they comprehend these early gestures when they are produced by another person; if they do not, they cannot be viewed as bidirectional communicative conventions. The same may even be true of some of children's earliest "words," which may be learned through ritualization as well. That is, children may have some early "words" that they have acquired by simply mimicking an adult sound and then the adult responds in some predictable and interesting way. These so-called pre-symbolic forms are often characterized as being simply a part of an activity, not a symbol standing for anything else in the activity (Bates 1979). If they are of this nature, and they are not truly conventional (the child does not understand the adult's understanding of them), they are best called vocal signals.

The argument is thus that the child's earliest words depend in fundamental ways on her ability to perceive and understand the actions of other persons as intentional, and especially her ability to understand that other persons act intentionally towards her intentional states. It could be argued, of course, that this account does not solve the mystery of referential indeterminacy, but simply shifts it to a new domain: how does the child understand the intentions of others when they can be interpreted in many ways as well (Levinson 1995)? There are a number of different views of this

problem, but my own view is that the process must get started by the infant recognizing in other persons intentions that are in some sense “the same” as intentions that she has experienced herself in her own intentional actions (Tomasello 1995b). In any case, even if the process of intention reading is at this point as mysterious as the process of linguistic reference that it is supposed to solve, at least the mystery is localized appropriately, in my view, so that future investigators will know where to look for more definitive answers. I should add that although some constraints theorists have discussed the problem of how children initially come to understand linguistic reference (e.g. Markman 1989; see also the Golinkoff *et al.* 1993 proposal for word learning principles), none has provided a satisfactory account beyond the proposal of Macnamara (1982) that reference is an ontogenetic “given.”

A final question is whether, in addition to social-pragmatic information, children need word learning constraints (Woodward & Markman 1998). It could reasonably be argued that the child’s understanding of adult communicative intentions presupposes a bias in making certain ontological distinctions – for example, a bias towards whole objects when the child understands the adult to be directing her attention to a doggie. But we must be very clear that what distinguishes the constraints position is not the claim that children are biased to conceptualize the world in certain ways – that is a claim that no theory of intersubjective communication can do without – but the claim that certain conceptualizations are tied to language in specific ways *a priori*. The social-pragmatic approach recognizes explicitly that the process of word learning depends fundamentally on the child being biased to conceptualize the world in certain ways (similar to adults’ conceptualizations), it is just that the connection of conceptualizations to language must be learned in communicative interactions with others.

4 Conclusion

My argument may be summarized in an ontogenetically forward direction as follows. In the first half of the second year of life young children begin to learn language. The major social cognitive skill that underlies their ability to do this is their understanding of the intentional actions of others, especially their understanding that other persons have intentions towards their intentional states. Children’s ability to reproduce these intentional communicative actions via some form of cultural or imitative learning involves a role reversal – the child has intentions towards the other person’s intentional states – which leads to the creation of linguistic conventions. Linguistic conventions are most clearly distinguished from other forms of communication precisely by virtue of their bidirectional or “shared”

nature. Then, in the second half of the second year of life, young children go on to become more and more skilled at determining precisely what are the referential intentions of others in particular situations. They learn to do this in all kinds of ongoing social interactions, whose complexity and diversity preclude the possibility that this learning is of a straightforward mechanical variety in which word is mapped to world – either with or without *a priori* constraints.

Perhaps due in part to the attention paid to children's so-called "theories of mind" at four years of age (involving their understanding of the beliefs of others), the importance of infant understanding of other persons' intentions as a social cognitive skill is only now being fully realized. Its great importance can perhaps best be appreciated when human children are compared with other organisms who do not possess this skill. Thus, the naturally occurring gestures of our nearest primate relative, the chimpanzee, resemble human gestures and words in a number of ways, but they are also different in a number of important ways as well. Most importantly for current purposes, chimpanzees in their natural habitats do not seem to acquire their skills of gestural communication via any form of cultural or imitative learning, but rather by ritualization, and so they do not develop any communicative conventions that are bidirectionally shared (Tomasello & Camaioni, 1997). There are various lines of evidence that chimpanzees do not understand others as intentional agents and that this is why they do not form such conventions, although chimpanzees raised and treated intentionally by humans may move some distance in this direction (Tomasello 1998). The same general point may be made about severely autistic children who, because of their biological deficit, do not seem to be skilled at understanding others as intentional agents with whom they may share attention, and so many of them learn little conventional language as a result (Landry & Loveland 1986).

Children's amazing skills at determining the specific referential intentions of adults in specific communicative circumstances are only now being fully appreciated. Indeed, in addition to the many and varied word learning contexts represented by the experimental studies summarized in this chapter, there is also observational evidence that children in some, perhaps many, of the world's cultures even tune into the flow of social interaction between third parties and learn some linguistic conventions from *their* linguistic interactions (Brown and de León, chs. 17 and 18, this volume). Acquisition in this circumstance also depends on children's ability to perceive and understand the communicative intentions of others, but the specific social information used may be different. And we should not neglect, in all of this concern with understanding nonlinguistic intentions, that, as children learn more language, language itself provides additional

information for understanding the intentions of others. For example, as children learn a number of words in a particular semantic domain, acquiring new words in that domain becomes easier as a new word's communicative significance may be contrasted with that of the known words – perhaps on the pragmatic inference that people use new terms for new communicative intentions (Clark 1988). In addition, as children become grammatically more sophisticated, the linguistic context within which they hear new words becomes an increasingly important source of information as well (Brown 1973; Gleitman 1990).

Although there is no room to make the argument fully here, it is also important that children's later development in the more grammatical aspects of language may also be seen in these same intentional terms if grammatical competence is viewed in functional terms. In this case the process may be seen again as adults using linguistic forms with communicative functions, which children then appropriate for their own use by means of some form of cultural or imitative learning (Tomasello & Brooks 1999). It is just that in this case the linguistic forms they are learning are relatively abstract (e.g. tense markers), or else they are whole linguistic constructions, many of which are only categorically and schematically specified. The communicative function of linguistic constructions that guides the learning process is the symbolization of whole events with their participants, providing specific perspectives on those events and participants given the particularities of the current discourse context (Clark 1990). This process is especially clear in children's earliest linguistic constructions, which remain tied very concretely to specific events and the words used to talk about them (Tomasello 1992a).

To conclude, I will simply reiterate my view that the cognitive foundations for language acquisition are two: (1) children's growing ability to conceptualize the world in something like the same way as adults, and (2) children's growing ability to understand adults' communicative intentions towards particular aspects of that world in particular communicative circumstances. Future research on the cognitive bases of language acquisition should pursue the ontogeny of both of these cognitive foundations with equal vigor.

NOTE

Thanks to Steve Levinson and Melissa Bowerman for very helpful comments on an earlier version of this chapter.

REFERENCES

- Akhtar, N., M. Carpenter, & M. Tomasello. 1996. The role of discourse novelty in children's early word learning. *Child Development* 67: 635–645.

- Akhtar, N., & M. Tomasello. 1996. Twenty-four-month-old children learn words for absent objects and actions. *British Journal of Developmental Psychology* 14: 79–93.
1999. Intersubjectivity in early language learning and use. In S. Braaten (ed.), *Intersubjective communication and emotion in ontogeny*. Cambridge: Cambridge University Press, 316–335.
- Baldwin, D. A. 1991. Infants' contribution to the achievement of joint reference. *Child Development* 62: 875–890.
- 1993a. Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language* 20: 395–418.
- 1993b. Early referential understanding: young children's ability to recognize referential acts for what they are. *Developmental Psychology* 29: 1–12.
- Bates, E. 1979. *The emergence of symbols: cognition and communication in infancy*. New York: Academic Press.
- Bloom, L. 1993. *The transition from infancy to language: acquiring the power of expression*. Cambridge: Cambridge University Press.
- Brown, R. 1973. *A first language: the early stages*. Cambridge, MA: Harvard University Press.
- Bruner, J. S. 1983. *Child's talk: learning to use language*. New York: Norton.
- Carpenter, M., N. Akhtar, & M. Tomasello. 1998. Sixteen-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development* 21: 315–320.
- Clark, E. 1988. On the logic of contrast. *Journal of Child Language* 15: 317–336.
1990. Speaker perspective in language acquisition. *Linguistics* 28: 1201–1220.
- Dunham, P. J., F. S. Dunham, & A. Curwin. 1993. Joint attentional states and lexical acquisition at 18 months. *Developmental Psychology* 29: 827–831.
- Fillmore, C. 1982. Frame semantics. In Linguistic Society of Korea (ed.), *Linguistics in the morning calm*. Seoul: Hanshin, 111–138.
- Gergely, G., Z. Nádasdy, G. Csibra, & S. Biró. 1995. Taking the intentional stance at 12 months of age. *Cognition* 56: 165–193.
- Gibson, E., & N. Rader. 1979. Attention: the perceiver as performer. In G. Hale & M. Lewis (eds.), *Attention and cognitive development*. New York: Plenum Press, 6–36.
- Gleitman, L. R. 1990. The structural sources of verb meanings. *Language Acquisition* 1: 3–55.
- Golinkoff, R., K. Hirsh-Pasek, C. B. Mervis, W. Frawley, & M. Parillo. 1995. Lexical principles can be extended to the acquisition of verbs. In M. Tomasello & W. E. Merriman (eds.), *Beyond names for things: young children's acquisition of verbs*. Hillsdale, NJ/Hove: Lawrence Erlbaum, 185–221.
- Golinkoff, R. M., C. B. Mervis, & K. Hirsh-Pasek. 1994. Early object labels: the case for a developmental lexical principles framework. *Journal of Child Language* 21: 125–155.
- Lakoff, G. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- Landry, S., & K. Loveland. 1986. Joint attention in autism and developmental language delay. *Journal of Autism and Developmental Disorders* 16: 335–349.
- Langacker, R. W. 1986. An introduction to cognitive grammar. *Cognitive Science* 10: 1–40.

1987. *Foundations of cognitive grammar*, vol. 1. Stanford, CA: Stanford University Press.
1991. *Foundations of cognitive grammar*, vol. 2. Stanford, CA: Stanford University Press.
- Levinson, S. 1995. Interactional biases in human thinking. In E. Goody (ed.), *Social intelligence and interaction*. Cambridge: Cambridge University Press, 221–260.
- Macnamara, J. 1982. *Names for things: a study of human learning*. Cambridge, MA: MIT Press.
- Markman, E. M. 1989. *Categorization and naming in children: problems of induction*. Cambridge, MA: Bradford/MIT Press.
1992. Constraints on word learning: speculations about their nature, origins, and word specificity. In M. Gunnar & M. Maratsos (eds.), *Modularity and constraints in language and cognition*. Hillsdale, NJ: Lawrence Erlbaum, 59–102.
- Meltzoff, A. 1988. Infant imitation after a 1-week delay: long term memory for novel acts and multiple stimuli. *Developmental Psychology* 24: 470–476.
1995. Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental Psychology* 31: 838–850.
- Nelson, K. 1985. *Making sense: the acquisition of shared meaning*. New York: Academic Press.
- Quine, W. V. O. 1960. *Word and object*. Cambridge, MA: MIT Press.
- Saussure, F. de. 1916. *Course in general linguistics*. New York: Philosophical Library.
- Sperber, D., & D. Wilson. 1986. *Relevance: communication and cognition*. Cambridge, MA: Harvard University Press.
- Talmy, L. 1996. The windowing of attention in language. In M. Shibatani & S. Thompson (eds.), *Grammatical constructions: their form and meaning*. Oxford: Oxford University Press, 235–238.
- Tomasello, M. 1992a. *First verbs: a case study of early grammatical development*. Cambridge: Cambridge University Press.
- 1992b. The social bases of language acquisition. *Social Development* 1: 67–87.
- 1995a. Pragmatic contexts for early verb learning. In M. Tomasello & W. E. Merriman (eds.), *Beyond names for things: young children's acquisition of verbs*. Hillsdale, NJ/Hove: Lawrence Erlbaum, 115–146.
- 1995b. Joint attention as social cognition. In C. Moore & P. Dunham (eds.), *Joint attention: its origins and role in development*. Hillsdale, NJ/Hove: Lawrence Erlbaum, 103–141.
1996. The cultural roots of language. In B. Velichkovsky & D. Rumbaugh (eds.), *Communicating meaning: the evolution and development of language*. Mahwah, NJ: Lawrence Erlbaum.
1998. Social cognition and the evolution of culture. In J. Langer & M. Killen (eds.), *Piaget, evolution, and development*. Mahwah, NJ: Lawrence Erlbaum, 221–246.
- Tomasello, M., & N. Akhtar. 1995. Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development* 10: 201–224.
- Tomasello, M., & M. Barton. 1994. Learning words in non-ostensive contexts. *Developmental Psychology* 30: 639–650.
- Tomasello, M., & P. J. Brooks. 1999. Early syntactic development: a construction grammar approach. In M. Barrett (ed.), *The development of language*. Hove: Psychology Press, 161–190.

- Tomasello, M., & L. Camaioni. 1997. A comparison of the gestural communication of apes and human infants. *Human Development* 40: 7–24.
- Tomasello, M., & J. Farrar. 1986. Joint attention and early language. *Child Development* 57: 1454–1463.
- Tomasello, M., & A. Kruger. 1992. Joint attention on actions: acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language* 19: 311–334.
- Tomasello, M., A. C. Kruger, & H. H. Ratner. 1993. Cultural learning. *Behavioral and Brain Sciences* 16: 495–552.
- Tomasello, M., R. Strosberg, & N. Akhtar. 1995. Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language* 22: 1–20.
- Wittgenstein, L. 1953. *Philosophical investigations*. New York: MacMillan.
- Woodward, A. L., & E. M. Markman. 1998. Early word learning. In W. Damon, D. Kuhn, & R. S. Siegler (eds.), *Handbook of child psychology*, vol. 2: *Cognition, perception, and language*. New York: John Wiley, 371–420.