# Links between Social Understanding and Early Word Learning: Challenges to Current Accounts

Dare A. Baldwin & Louis J. Moses, *University of Oregon*

**Abstract**

*If young children approached word learning with little social savvy, certain predictable patterns of error would arise in the way they interpret new words. The absence of such errors provides evidence that social understanding informs word learning even in the infancy period. We outline such evidence, and then scrutinize it with respect to four challenges. 1) Is it necessary to invoke genuine social understanding to explain infants' word-learning successes? 2) Do infants treat social clues as* criterial *in their interpretation of new words? 3) Individuals suffering clear deficits in social understanding sometimes display apparently intact vocabulary acquisition: Must we then conclude that word learning can proceed without the aid of social understanding? 4) Is processing of social clues too effortful to be generally useful for everyday word learning? The first challenge is answered by the available evidence: Infants indeed capitalize on social understanding to interpret new words. Although the remaining challenges have yet to be resolved, we offer speculations that might profitably guide future investigation.*

*Keywords*: Social Understanding; Communicative Intentions; Semantic Acquisition; Word Learning

Happily playing with novel toys while we watched the news, our young son—just turned two—suddenly announced 'No legal precedent!' Initially paralyzed by such remarkable command of legal jargon in one so near and dear, we soon realized that, of course, he was simply echoing what he'd heard emanating from the television. Parental gullibility and precocious verbal imitation aside, this anecdote genuinely captures something enduringly impressive. It illustrates one way in which children put social understanding to work for language learning. To explain: Among the crucial skills for language learning is the ability to link words with relevant objects. Research documents that children establish such links most readily when they hear words right at the time they are attending to the correct objects (e.g., Tomasello & Farrar, 1986). In the scenario we witnessed, these conditions were met: Our son was hearing language—from the television—just when he was focusing on objects of great interest to him. His arresting vocalization confirms that he was processing the

Address correspondence to Dr. Dare Baldwin, Department of Psychology, 1227 University of Oregon, Eugene, OR 97403-1227, USA, baldwin@darkwing.uoregon.edu.

© Blackwell Publishers Ltd. 2001. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.

language. Nevertheless, he apparently resisted interpreting that language as relevant to the objects at the center of his attentional focus. That is, quite appropriately, he showed no subsequent sign of thinking 'legal' or 'precedent' was the name for any of these objects. This is fortunate. If he were inclined to faithfully register links between word and object simply because they coincided, errors would be a frequent occurrence for him.

The point is this: The ability to *disregard* potential links between words and things in the world is as valuable for language learning as the ability to register such links, provided that one is skilled at recognizing which situations call for which strategy. And social understanding provides the key to this crucial distinction. The best evidence on which to base the formation of new links between words and things in the world is social in kind: evidence that speakers are intending to talk about those things. Conversely, it is usually safest to disregard potential links if speakers show no signs of an intention to refer to the objects at hand.

Is it plausible to think that young children might appreciate the relevance of clues to other's intentions—we will call these 'social clues' for brevity's sake—for guiding their learning of new words? A rich and diverse set of such clues (including gaze direction, head direction, body posture, voice direction, gestures, intonational quality, facial expression, and the like) is of course readily available to children in the everyday learning situation. In question, however, is whether children actively capitalize on the available social clues to assist them in drawing inferences about new words.

A substantial body of evidence now indicates that, as early as 12–18 months, children already approach language learning from just such a social stance. Our goal in this paper is to provide a brief overview of this evidence, and then to question it. Some challenges to the social understanding account have already been raised. We will describe these challenges, and evaluate where things stand. We will also offer some new challenges. In some cases, existing evidence from other areas of research provides relevant, but we think inconclusive, information to resolve these challenges. The product of all this? We aim to identify avenues for new investigation that might clarify social processes fundamental to language learning.

## Social Understanding Utilized for Language Learning: Evidence

The specific skill our son displayed in front of the television—appropriately resisting potential, but invalid, links between words and objects—has been documented in infants as young as 19–20 months. For example, in two studies Baldwin and colleagues (Baldwin, Markman, Bill, Desjardins, Irwin, & Tidball, 1996) showed infants some novel objects, and when they were focused on one of those objects, infants heard a novel label (e.g., 'A *dawnoo*!') produced with all the special intonational qualities typical of speech to infants. In one case, the speaker was seated next to infants, gazing toward the object at the time she produced the novel label. Infants readily learned the new word: When later asked to 'Find the *dawnoo*!', for instance, they systematically selected the correct object rather than an equally salient distracter object. A second, more noteworthy, condition provided direct evidence that infants relied on social clues to direct their new word-object mappings. This second case was equivalent to the first condition except that the speaker was seated out of infants' view. Hence she produced no evidence of intending to label the object on which infants were focused. From an adult's perspective, her voice emanated from the 'irrelevant beyond.' Infants apparently shared adults' experience of this situation. Despite the contiguity of novel label

© Blackwell Publishers Ltd. 2001 

and novel object, infants resisted registering a new word-object link. When probed for comprehension of the new label, they showed utterly unsystematic performance, selecting indiscriminately between the 'correct' and distracter objects.

Based on these findings, Baldwin et al. concluded that infants as young as 19–20 months rely on a social criterion for establishing links between words and objects. That is, infants seem to view temporal contiguity alone as inadequate grounds on which to establish a new word-object link; instead, infants require evidence that the speaker is intending to talk about the object at issue. They bring their social understanding to bear in deciding when word-object co-occurrence is noteworthy.

A large body of additional evidence supports this social understanding account. Details of the relevant studies have been reviewed in several places (see, for example, Baldwin & Tomasello, 1999; Bloom, 2000; Tomasello, 1992, 1999); for this reason, we will simply outline what has been concluded in each case regarding infants' ability to recruit early social skill to aid word learning.

Baldwin's prior research (1991, 1993a) had already provided evidence that infants as young as 18 months actively gather social information to guide their inferences about word meanings. In those studies, infants were presented with a potentially tricky—although altogether common—word-learning situation: They were looking at a novel object just when the speaker was looking at and labeling (saying, for example, 'A *toma*!') a completely different novel object. Infants actively gathered social information to help them negotiate the complexities of this 'discrepant labeling' situation. On hearing the novel label, they quickly checked the speaker's face, noted her gaze direction, and appreciated the significance of these clues for word learning. When probed for comprehension of the novel label, infants systematically selected the object the speaker had been gazing toward at the time of labeling, rather than the object they themselves had been focused on at the time the label was heard (see Dunham, Dunham, & Curwin, 1993, for corroborating evidence).

One question left unanswered by the initial research was precisely why infants bothered to check the speaker's face on hearing novel labels. Baldwin (1991, 1993a) found that they did so more frequently when the speaker's attentional focus was discrepant from infants' at the time of labeling, suggesting that infants may have been aware of a possible discrepancy in focus and so checked the experimenter's gaze specifically to clarify her reference. If this were correct, such face checking would, in and of itself, indicate social understanding put to work for word learning. It seemed possible, however, that infants checked the speaker's face on hearing novel labels merely as a result of simple orienting processes. After all, speech is noisy, and infants may have oriented to the sound.

Baldwin, Bill, and Ontai (1996) developed a procedure specifically designed to test between the social understanding and simple orienting accounts. They presented infants of 12 and 18 months with one and the same orienting stimulus—novel labels produced without accompanying motion—in the presence of either one novel object versus two novel objects. If orienting alone were the trigger for infants' face checking, then their checking should be unaffected by the number of objects present. Alternatively, if social understanding underlay face checking, then infants should show increased checking when labels were provided in the presence of two objects relative to just one object. The speaker's intended reference would be more ambiguous in the former instance, provoking infants to gather social clues for clarification. Indeed, infants of both 12 and 18 months displayed significantly more face checking when reference was ambiguous (the case of two novel objects). Moreover, baseline levels of

© Blackwell Publishers Ltd. 2001

face checking in the absence of novel labels did not differ for one versus two novel objects, clarifying that the difference associated with novel labels was not simply the result of a general increase in checking rates with two objects. These findings thus buttress the viability of a social understanding explanation for infants' face checking, and indicate that infants even as young as 12 months are already actively using information in the speaker's face to guide inferences about word meaning.

Flexibility is something infants also exhibit when exploiting social clues for word learning. In the studies described thus far, a collection of clues including gaze direction, body posture, voice direction, and hand position all provided information about the speaker's intended referent. However, in some situations any or all of these particular clues may be uninformative. Are infants then at a loss? Or do they fall back on simpler strategies such as attending to simple associative information? Neither. Ingenious research by Akhtar, Tomasello and colleagues documents that infants at least as young as 18 months adaptively utilize alternate clues to meaning in such cases (Akhtar & Tomasello, 1996; Tomasello & Barton, 1994; Tomasello, Strosberg, & Akhtar, 1996). They have shown, in particular, that infants can use speakers' emotional expressions and other action complexes (such as termination of search) as alternative sources of information about their intended reference when gaze or other simpler action clues are absent or uninformative. In two such studies (Tomasello & Barton, 1994; Tomasello, Strosberg, & Akhtar, 1996), for example, a speaker announced her intention, saying 'Let's go find the *toma*!', then displayed disappointment on retrieving a toy from a bucket. Later, on finding a second toy, she revealed pleasure, and terminated her search. Only her facial expressions and termination of search provided information about the intended referent; all other actions, such as gaze direction, body posture, voice direction, and the like were equivalent across objects. Yet, as indexed by their subsequent comprehension and production performance, infants as young as 18 months readily appreciated that the object eliciting pleasure and search termination in the speaker was the 'toma.' Importantly, infants succeeded despite the fact that they did not even see the correct object until some time after the label was produced, and actually saw a different, distracter, object in the intervening period. Apparently, infants give priority to social clues in determining appropriate word-object links, and adaptively capitalize on whatever social clues are available in a given learning situation (work by Baldwin, 1993b, further documented these skills).

One might wonder whether the abilities infants displayed in the foregoing studies are tightly bound to word learning, or even more narrowly, to the learning of object labels. That is, are these skills part of a pre-adapted package dedicated to count noun acquisition? Clearly not. Infants recruit these skills for verb learning as well (Akhtar & Tomasello, 1996; Tomasello & Barton, 1994). Moreover, they can utilize such skills for purposes other than word learning, such as to interpret others' emotional messages (Moses, Baldwin, Rosicky, & Tidball, in press). For example, on hearing an adult's outburst of disgust ('Iiuu! Yecch!') infants of both 12 and 18 months immediately checked the adult's face, and subsequently remained wary of the object the adult was looking at when the outburst was provided. At the same time, infants remained relatively sanguine about a distracter object they themselves had been focused on at the time the outburst was actually heard. Thus infants spontaneously consulted the speaker for social clues to clarify the reference of an emotional message, and used those clues to guide their own behavior toward the relevant objects.

More recent research sheds light on the scope and development of young children's ability to deploy social understanding to aid language learning. We offer just a few

© Blackwell Publishers Ltd. 2001

examples. Akhtar and Callanan (in press), for instance, have shown that two-year-olds can interpret the meaning of new words even when those words were actually addressed to someone else. By age two, then, and quite possibly earlier, children take advantage of the language available within their broader social setting, and they put their 'eavesdropping' skills to work for language learning.

As children move into the preschool period they refine their social inference skills, and these refinements promote word learning in new ways. Preschool-aged children are increasingly acute observers of others' actions, and gain additional proficiency at discerning subtleties of intention beneficial to inferences about word meaning. For instance, Saylor, Sabbagh, and Baldwin (under revision) found that 3- and 4-year-olds could process a series of novel gestures to infer the speaker's intention to mark a contrast in meaning. In the critical study, Saylor et al. presented preschoolers with a familiar object, such as a spider. The spider was purple with just one small part in a different color (e.g., a small portion of the abdomen was pink). In the baseline, *no contrast* condition, the experimenter showed children the picture, saying 'See this?' and then provided what turned out to be an ambiguous point—pointing on high in the general direction of the part—while saying 'What color is the *pedicel*?' Preschoolers were at a loss. They answered 'pink' or 'purple' indiscriminately, suggesting they were uncertain as to whether the new word referred to the whole object (the spider), or the part (the pedicel). However, when the identical ambiguous point contrasted with a previous gesture that clearly indicated the whole object, children readily interpreted the otherwise ambiguous point as indicating the part, answering 'pink' at high rates. That is, in the *gesture contrast* condition, the experimenter first provided a 'whole-object' gesture—circling the whole object with index finger touching the page while saying, 'See this?' She then followed up with the very same ambiguous point that had been used in the *no contrast* condition—pointing toward the object from on high—while asking 'What color is the *pedicel*?' What is interesting here is that children's increased ability to interpret the new word as referring to the part in the gesture contrast condition could only have been achieved through appreciation of the *contrastive* force of the two gestures. We argue this for two reasons: First, their unsystematic performance in the *no contrast* condition showed that the pointing gesture that accompanied the novel word did not clarify anything, and the identical point was used across both conditions. Second, all that was added in the *gesture contrast* condition was the gesture encircling the *whole object* which, if interpreted in isolation, would only serve to enhance children's attention to the whole object, not the part. Thus children must have appreciated that the prior whole object gesture, by virtue of contrast, clarified the reference of the otherwise ambiguous point. Also noteworthy is that children probably had not previously experienced circling versus pointing gestures being used to mark an intention to contrast the meaning of two terms. Nevertheless, they readily interpreted the contrastive intention in this novel format, and used it to infer appropriate meanings that were otherwise opaque to them.

In addition to increasingly sophisticated inferences about speakers' intentions, preschoolers begin to be sensitive to clues linked to speakers' *epistemic* states, such as their knowledge or beliefs, and use these clues to direct language learning. Documenting this, Sabbagh and Baldwin (in press) found that 3- and 4-year-olds resisted learning a new word if the speaker displayed signs of ignorance (e.g., the speaker said, 'I'm not sure . . . maybe this one is the *blicket*' while pointing and looking at the target object), whereas they readily learned the new word when the speaker seemed knowledgeable (e.g., the speaker said, 'I know . . . this one is the *blicket*').

© Blackwell Publishers Ltd. 2001 

All in all, recent research indicates that infants and preschoolers readily take advantage of social information available in the context of ongoing, dynamic interaction to guide learning across a range of domains. Their ability to do so is greatly to the benefit of language learning, as well as knowledge acquisition more generally (Baldwin, 2000; Baldwin & Moses, 1996; Bloom, 2000), and may also lie at the heart of humans' special facility for complex cultural participation (Tomasello, 1999).

## Challenges

Central to the social understanding account is the specific claim that even infants readily process others' actions for clues to their intentions—intentions to talk about things, find things, communicate emotions regarding things or experiences, and the like. The appropriateness of this 'high-level' interpretation is open to question on several counts. For one, is it necessary to attribute genuine understanding of intentions to infants to account for their success in existing studies? We feel confident in doing so, and outline our reasoning below. Even with reassurance on this point, a range of additional questions arises about the social understanding account, and we will consider these in turn.

*Social understanding versus processes of social influence*. Samuelson and Smith (1998) recently challenged the social understanding account, arguing that attributing social *understanding* to infants based on existing findings was too generous. Instead, they suggested that the findings can be explained more parsimoniously in terms of fundamental cognitive processes, such as memory and attention. In particular, they proposed that social interaction affects what children attend to, and how motivated they are to recall things. Put another way, they proposed that processes of social *influence* might directly govern children's word-referent links by shaping attention and memory; social understanding might play little or no role.

Samuelson and Smith articulated their challenge through consideration of a plausible alternative explanation for findings presented by Akhtar, Carpenter, and Tomasello (1996). Akhtar et al. had interpreted these findings as indicating a previously unrecognized level of sophisticated social understanding on infants' part. Briefly, in the study of particular interest, Akhtar et al. presented 24-month-olds with three novel objects. Infants examined two of these novel objects in the company of two experimenters; then one of the experimenters departed while the remaining experimenter presented the third novel object. When the absent experimenter returned, she looked at the array of three novel objects (one of which was novel to her, given her absence during its presentation), and said 'Look! I see a *gazzer*! A *gazzer*! I see a *gazzer* in there!' She was standing at enough of a distance from the array that her gaze did not clarify which of the three objects she intended to refer to, and she provided no clarifying gestures. A subsequent comprehension and production test revealed that infants reliably linked the new label, *gazzer*, with the object that had been novel to the speaker. Notably, the relevant object was not novel to infants themselves; hence they must have recognized the significance of its novelty from the perspective of the speaker. Akhtar et al. concluded from their findings that infants appreciate the role of discourse novelty as a clue to the intentions underlying others' messages, and use such understanding to drive their inferences about the meanings of new words.

Samuelson and Smith pointed out, however, that processes of social influence—what they termed 'contextual processes'—could account for the findings without any need to attribute an understanding of discourse novelty to infants. Specifically, the

© Blackwell Publishers Ltd. 2001

object that was novel to the speaker, though not novel to infants, nevertheless gained greater salience within the social context by virtue of being treated distinctively (e.g., it was presented later, in the presence of only one experimenter, etc.) relative to the other two novel objects. The target object's special treatment may thus have garnered it greater attention from infants. If so, attentional enhancement alone could account for infants' tendency to link the new label with the object novel to the speaker, rather than any appreciation for such novelty from the speaker's perspective.

As it happened, Akhtar et al. had incorporated a control condition to rule out such a low-level interpretation. In this control, the lead-in scenario was identical, in all but one respect: the target object received special treatment, as described earlier, but the experimenter simply said 'Look at that!' rather than 'Look at the *toma*!' Infants in the control condition subsequently were asked comprehension questions regarding the array of three novel objects that were identical to those asked in the experimental condition. If attentional salience were the *sole* reason for infants' selection of the target object in response to comprehension questions in the experimental condition, then they should show equivalently high levels of target object selection in response to comprehension questions in the control condition. However, it turned out that infants in the control condition rarely selected the target object in response to comprehension questions, a pattern quite different from what they showed in the experimental condition. Attentional factors alone did not seem to be driving infants' inferences about word meaning.

Still, it remains conceivable that attentional salience could have been strong enough to direct infants to link the novel label with the object given special treatment, but only when the label was supplied in the learning phase, as it was in Akhtar et al.'s experimental condition. Attentional salience may not have been strong enough to persist all the way through a series of comprehension questions (which it would have had to do to influence comprehension performance in the control condition). If so, comprehension performance in the experimental condition would show systematic selection of the target object, while comprehension performance in the control condition would not, and this is just the pattern Akhtar et al. found. For this reason, the Akhtar et al. control condition may not entirely rule out Samuelson and Smith's social influence alternative explanation.

Samuelson and Smith undertook to directly investigate their alternative account. They constructed a word-learning scenario with just one crucial difference from that of Akhtar et al.: One of three novel objects received special treatment, but discourse novelty was not involved. In the Samuelson and Smith scenario, two experimenters presented two novel objects to infants in one location (the floor), and then moved to a new location (a table draped with a special tablecloth) to present a third novel object. On returning to the original location, all three objects were arrayed together, and one experimenter produced the novel label (e.g., 'There is a *gazzer* in here. There is a *gazzer*. Look there is a *gazzer* . . .'). To subsequent comprehension/production questions, infants tended to select the object that had received special treatment—in this case, presentation in a unique location—as the referent for the label *gazzer*.

Samuelson and Smith included a control condition (also modeled on that of Akhtar et al.) identical to their experimental condition except that the experimenter provided no novel label, simply saying 'Look in here! There it is! . . .' When subsequently tested for comprehension of the novel label, infants seldom selected the target object. The different pattern of responding in the control versus experimental conditions was important, in that it demonstrated that infants' response to comprehension questions

© Blackwell Publishers Ltd. 2001

in the experimental condition reflected genuine word learning, rather than simple preferential responding.

Clearly, there is no need to attribute an appreciation for discourse novelty to infants to account for their performance in Samuelson and Smith's word-learning scenario, because discourse novelty was not involved. Special treatment alone seemed sufficient to direct infants' word learning. On the face of it, then, these findings seem to support the viability of the social influence account at the expense of the social understanding account. But caution is warranted. In fact, the Samuelson and Smith findings are as ambiguous with respect to these accounts as were the original Akhtar et al. results. While it is possible that special treatment alone determined the new word-object link infants established in the Samuelson and Smith experimental condition, it is also entirely possible that processes of social understanding—although not specifically an understanding of discourse novelty—helped to guide this link. That is, infants may have inferred that the speaker in the Samuelson and Smith study was intending to refer to the object to which she'd given special treatment; after all, the situation made abundantly clear that she favored that object over the other two, or at least regarded it as distinctive. Such signs of special regard could well have been a basis for inferring that she intended to refer to that object when producing the novel label.

Thus the Samuelson and Smith findings actually provide no guidance as to whether social influence, social understanding, or both were at work in children's word learning. Clearly, they do not *rule out* a role for social understanding in word learning; they merely leave us wondering.

Other available evidence, however, can help us decide between the two competing accounts. Samuelson and Smith's challenge can be construed in at least two ways: a) in terms of reservations about the viability of attributing social understanding *at all* to infants, or b) in a more limited way, in terms of skepticism that such social understanding is readily employed in the aid of word learning. We argue that existing evidence effectively counters both ways of framing their challenge.

First, infants' functioning in other arenas, such as imitation, provides converging evidence that they are genuinely capable of inferring others' intentions as early as 15–18 months (Meltzoff, 1995, Meltzoff & Brooks, in press). For example, Meltzoff demonstrated that 15- and 18-month-old infants readily reenact a novel action sequence to achieve a goal after having witnessed an adult repeatedly attempt, but fail, to achieve the same goal. Thus infants imitate not the surface pattern of another's action itself, but rather the action that would satisfy the intention they infer from viewing such action. Likewise, Carpenter, Akhtar, and Tomasello (1998) found that infants are inclined to imitate motions that are produced intentionally rather than accidentally, despite the equivalent salience of those actions. Clearly, then, infants process people's motions to glean information about underlying intentions, and base their responses to such motions on these social inferences.

Second, infants do indeed put inferences about intentions to work in the specific case of word learning. Recognizing that processes of social influence of the kind emphasized by Samuelson and Smith might well play an important role in children's word learning, many researchers approached their investigation of the social understanding account in terms of whether social understanding might operate as well, *over and above* such social influence mechanisms. Thus the bulk of studies described in our earlier section actually included controls specifically designed to rule out processes of social influence as the sole determining force in infants' word learning, in order to showcase the role of social understanding (e.g., Baldwin, 1991, 1993a, 1993b,

© Blackwell Publishers Ltd. 2001

Baldwin, et al., 1996; Tomasello & Barton, 1994; Akhtar & Tomasello, 1996). These controls targeted, in particular, the possibility that socially-induced salience of the target object might have accounted for children's tendency to link the new word with the target object. For example, even in Baldwin's (1991, 1993a) first research on these issues, infants could succeed in exhibiting social understanding only if they inhibited a link between the novel word and the salient object on which they were focused at the time of labeling. Processes of social influence, as conceptualized by Samuelson and Smith, would have led infants to link the new word with the object salient to them at the time of labeling, rendering them prone to a word learning error. As it turned out, they did not make such errors. Some additional evidence against the social influence account arises from the finding across both studies that infants in a preference-control condition showed quite a different pattern of responding from those who were tested for comprehension of the novel label. Rather than being asked comprehension questions after presentation of the novel label, infants in the preference-control condition were simply asked questions about object preferences, such as 'Point to your favorite one!' and 'Find the one you like!' If social influence processes increasing the salience of one object over another were at work alone, then infants in the preference control condition might be expected to display the same pattern of selection as did infants responding to comprehension questions. But they did not. Infants in the preference-control condition chose randomly. These findings further strengthen the case that social understanding was operating over and above social influence processes to drive infants' word learning.

Of course, as we described earlier with regard to a similar control incorporated by Akhtar et al., the preference-control findings do not, in and of themselves, provide an entirely definitive case against social influence processes. Other research does so, however. In a different set of studies, Baldwin (1993b) provided an additional source of evidence that attentional and memory factors alone do not drive infants' word learning. In a first study, infants were shown two opaque buckets, each of which was known to contain a novel object (but infants didn't know which object was in which bucket). The experimenter then produced a novel label (e.g., 'I see a *dawnoo*! It's a *dawnoo*!') in conjunction with action indicating an intent to refer (e.g., she raised the lid of one opaque bucket and peered in). Immediately after producing the novel label in the 'coincide' condition, the experimenter extracted the object hidden inside the acted-upon bucket and handed it to infants; after 15 seconds the other object was extracted and handed to them as well. Thus the social clues the experimenter exhibited during labeling coincided with the attentional salience linked to seeing an object immediately after labeling. In another condition, social clues were put into direct conflict with attentional salience. Specifically, in this 'conflict' condition, immediately after labeling the experimenter turned away from the bucket whose contents she'd just targeted with social clues, and instead extracted the object hidden inside the other, thus-far-ignored, bucket and handed it to infants. She did not give them the object toward which she'd directed social clues until at least 15 seconds later. In both conditions, then, infants began attending to a novel object immediately after hearing the novel label. In the coincide condition, the object seen first happened to be the very same object that the experimenter had shown signs of referring to when producing the novel label. In the conflict condition, however, the object referred to was not seen until at least 15 seconds after the other novel object. In this condition, then, clues to referential intent were forced to compete with powerful conflicting pressure from attentional salience. If mechanisms of salience alone determine infants' new word-object links, then they

© Blackwell Publishers Ltd. 2001 

should link the novel label with the object they saw first after labeling across both conditions. Alternatively, if an appreciation for referential intent informs their establishment of new word-object links, they should link the new label with the object seen first only in the coincide condition. In the conflict condition they should link the novel label with the object seen second, since this was the object toward which referential action had been directed at the time of labeling. Subsequent comprehension questions revealed the latter pattern, indicating that social understanding 'out-competed' attentional salience even when salience was quite powerfully instantiated.

Studies by Moore, Angelopoulos, and Bennett (1999) and Hollich, Hirsh-Pasek, and Golinkoff (2000) further strengthen the case that infants weight social clues more heavily than attentional salience when forming new word-object links. These researchers presented children with word learning situations in which salience was directly pitted against social clues. For example, Moore et al. looked at and labeled a novel object just at a time when a distractor novel object emitted flashing lights. Despite their patent interest in the distractor, children 24 months and older nevertheless linked the novel label with the object toward which social clues were directed. Similarly, Hollich et al. presented infants with objects differing in inherent salience; they found that infants 18 months and older linked a novel label with the object toward which a speaker directed social clues, even when this object was less salient relative to the distractor object. Clearly, infants favor social clues over salience to guide the formation of new word-object links.

Finally, and most conclusively, Diesendruck, Markson, Akhtar & Reuder (2001) recently demonstrated, in two studies, that salience-enhancing special treatment to a labelled object does not induce infants to establish a new word-object link, if the special treatment is unintentional on the part of the speaker. When intentional, in contrast, the self-same special treatment does engender a new mapping. These findings effectively resolve the important challenge raised by Samuelson and Smith.

Together, existing research provides abundant evidence that social understanding, over and above processes of social influence, is important to word learning. The joint functioning of these mechanisms is all to children's benefit, enabling them to capitalize effectively on the information-rich social context within which word learning occurs.

*Centrality of intentions in word learning*? While reassured that early social understanding operates to facilitate word learning, one might still question the centrality of such understanding in infants' inferences about word meaning. The evidence we've reviewed indicates that infants value social clues regarding others' intentions, and weight them more highly than some other factors—such as attentional salience—in their inferences about word reference. But perhaps infants go further and even treat social clues as *criterial* in establishing new word-to-world links. We first outline evidence suggesting that infants follow the more radical path. Then, we report on some recent conflicting findings and suggest possible empirical avenues for resolving the outstanding issues.

Baldwin and colleagues (Baldwin et al., 1996) provided two lines of evidence that a social *criterion* operates in infants' inferences about word reference. First, as reported earlier, they found that infants 19–20 months resist linking a novel label with a novel object at the focus of their attention if clues from the speaker are lacking to indicate that she is talking about that object. That is, when an out-of-view speaker produced the novel label, infants disregarded it as a possible name for the novel object they were focusing upon. In contrast, they readily linked label and object when the

© Blackwell Publishers Ltd. 2001 

label was produced by an in-view speaker gazing at that object. A manipulation embedded within one of these studies provided a second source of evidence that the speaker's intent is what matters to infants in their decisions about new word-object links. In one set of conditions, the speaker carefully produced the novel label just when infants shifted their gaze toward the target object; in the other set of conditions, the label was produced on a strict time-line, without reference to infants' focus of attention. If attentional salience at the time of labeling is important in the formation of new word-object links, then differences in the synchrony of labeling with infants' attentional focus on the target object should influence the likelihood of infants' establishing a new word-object link. Conversely, if signs of an intention to refer are of crucial importance to infants, then such differences in synchrony might not matter. And they did not. What made this finding especially noteworthy was that the synchrony manipulation did have a measurable and significant impact on infants' *interest* in the target object. That is, when infants heard labels synchronized with their focus on the target they gazed longer at the target than when labels were presented on a fixed schedule that often resulted in asynchronies with their focus on the target. Thus the synchrony manipulation enhanced the target object's salience, but did not increase the likelihood that infants would create a link between label and target object. In this collection of studies, then, the availability of social clues mattered to infants in their formation of new word-object links; in contrast, they seemed to regard salience as immaterial to word learning.

Another study (study 2 of Baldwin (1993b)) further substantiates the possibility that infants treat signs of referential intent as the key factor licensing the establishment of new word-object links. In this study, the experimenter produced *nonreferential* action (e.g., incidental, idle 'futzing' with the lid of one bucket) in conjunction with labeling. This nonreferential action enhanced infants' *attention* to the bucket to the same degree that referential action in a companion study described earlier (e.g., lifting the lid and peering into the bucket) had achieved. Despite equivalent impact on attentional salience at the time of labeling, referential versus nonreferential action did not have equivalent impact on infants' word learning. Referential action that accompanied labeling incited infants to establish a stable link between label and object; nonreferential action did not. These findings demonstrate, in a different manner, that infants specially key on clues regarding a speaker's intention to refer in establishing new word-object links; actions that highlight object salience in the absence of such intent don't seem to 'cut it' for infants.

In other research we (Moses, Baldwin, Rosicky, & Tidball, in press) found that infants operate according to a social criterion in the emotions domain as well as in the word-learning domain. We devised a methodology appropriate to the emotions domain that was directly analogous to the one, described earlier, that Baldwin, et al. (1996) had used for probing word learning. In our emotions research, 18- to 19-month-old infants were focused on a strange, ambiguous novel object when an experimenter produced an emotional outburst, either 'Oh! Nice!' or 'Iiuu! Yecch!' In one case, the experimenter producing the outburst was within infants' view and gazing toward the object of infants' focus. In the other case, the speaker was out of infants' view, and hence supplied no clues regarding the intended target of her emotional outburst. Infants hearing the outburst of disgust from the in-view experimenter studiously avoided approaching the target object. In contrast, infants who heard such disgust from the out-of-view experimenter were unaffected—they showed greater willingness to approach the target object. Apparently, infants discounted the disembodied outburst.

© Blackwell Publishers Ltd. 2001 

In the absence of social clues, they regarded it as uninformative despite the fact that it covaried with their focus on the target object.

Findings from several studies thus make a viable case that infants grant social clues criterial status in their inferences about word reference, as well as in their interpretation of emotional messages. Yet findings from several other recent studies seem to challenge this account. Although the goal of the conflicting studies was to investigate other phenomena, they happen to also document the formation of new word-object links in infants 14 months and up under learning conditions in which the novel labels were produced by a disembodied voice. That is, infants established links even though labels merely covaried with the relevant objects in the absence of signs of referential intent (Werker, Cohen, Lloyd, Casasola, & Stager, 1998; Stager & Werker, 1997; Schafer & Plunkett, 1998). For example, in the Werker et al. (1998) research, 14-month-olds were habituated to two phonologically dissimilar novel labels (e.g., *lif* and *neem*) emanating from a central speaker. Each label was associated with a distinctively different, brightly colored, moving, novel object on a monitor. In a subsequent test phase, infants looked longer at test trials in which change occurred in the pairings of label and object—change in the label, the object, or both. This pattern clarifies that infants had registered links between each label and the object it had been paired with during habituation. Such findings indicate the establishment of word-object links not licensed by signs of referential intent on the part of the speaker; they seem to directly conflict with those reported earlier from the Baldwin research. What can we make of such seemingly contradictory evidence?

Perhaps the criterial version of the social understanding account—for infants, only social clues license the formation of new word-object links—is incorrect, at least in its strongest form. Perhaps, after all, infants simply are especially attuned to social clues, and weight them heavily in cases of conflict with other sources of information, such as object salience. If object salience is not weighted as heavily as social clues in infants' word learning, its operation may be difficult to detect. As a 'weak' factor, null findings would be more likely when it is the only available information (as obtained in the studies that seemed to confirm the criterial version).

In fact, such openness to multiple sources of information might well be highly advantageous to infants, perhaps benefiting them even more than a rigid reliance on social clues as a definitive criterion for word learning. If infants are at all unsophisticated in their ability to read social clues—as seems likely—they might be impervious to subtle signs of intent that supply key information for word learning. Sensitivity to other sources of information would enable infants to learn words even when others' intentions remain opaque to them, provided that social clues typically are redundantly bundled together with social influence processes. And this, too, seems likely.

A contrasting, and more controversial, possibility retains the view that infants indeed treat social clues as criterial in determining relevant links between words and things in the world, and credits infants with new levels of subtlety in their inferences about intentions. In particular, perhaps infants are sensitive to subtle signs of social intent inherent in any given input situation, with repeated pairing of word and object serving as one form of evidence for such intent. For example, infants experiencing a strong association between a particular word and a particular object, no matter how disembodied the linguistic presentation, may construe the very fact of such extended covariation as itself evidence for social intent on the part of the disembodied speaker, provoking them to establish a new word-object link. That is, if a given word co-occurs with a particular object again and again across a few short minutes, this is likely no

© Blackwell Publishers Ltd. 2001 

accident, and infants, like adults, might be inclined to invoke social intentions to account for such cases. Consistent with this account, in the Werker et al. (1998) and Schafer and Plunkett (1998) research—in which word-object links emerged despite the absence of clues to referential intention from the speaker—infants heard the word in association with the target object many more times (i.e., 12 times and up) than in the Baldwin studies (i.e., 4 times or less).

How can we choose between these two accounts—privileged versus criterial status for social clues—so subtly but crucially different? The latter, social criterion, account can be put to the test. It predicts that infants should be inclined to construe extended word-world covariation as arising from the intentions of others and hence they should initiate word learning, all other things equal. But if evidence is supplied to infants indicating a different, non-social, explanation for such covariation, they should resist learning. For example, suppose infants gain independent experience of two distinct objects in sequence: 1) a toy clown emits a novel word—*zazoo*—each time a button is pressed, and 2) a novel contraption spits out a novel object each time a crank is turned. Suppose further that infants witness some books spontaneously falling off a shelf and bumping both the clown and the contraption, leaving their mechanisms jammed as a result, and *zazoo* just happens to be repeated again and again each time the novel object is ejected. If the criterial view is correct, infants might not, in this instance, take *zazoo* to be the name for the object. In contrast, they might readily learn this link after just the same number of co-occurrences if, in all cases, the correlation occurs by social design—that is, a person intentionally presses the button and turns the crank to produce *zazoo* when the novel object emerges. This study remains to be carried out.

As yet, then, we are left uncertain about the extent to which infants genuinely treat social clues as pivotal in their word learning. Also uncertain at present is the extent to which the ability to interpret social clues is *crucial* for word learning to proceed on a normal trajectory. That is, do deficits in reading social clues predispose children to developmental delays in vocabulary acquisition? Or can word learning proceed at a normal pace even in the absence of social understanding?

*Deficits in social understanding: Problematic for word learning*? We have argued that social understanding radically facilitates children's word learning: It enables them to avoid word learning errors (as when faced with discrepant labeling), and more broadly, it assists them in recognizing when words are relevant to things in the world, thus guiding the establishment of new word-to-world links. If this is correct, we might expect that deficits in social understanding would seriously undercut normal progress in word learning. Baron-Cohen and colleagues (Baron-Cohen, Baldwin, & Crowson, 1997) investigated this possibility. They examined word learning in children with autism, who are known to have deficits in social understanding as well as delays in language learning. In question was whether, by virtue of their social understanding deficits, children with autism would be especially prone to word-learning errors. They found that children with autism could readily link a new label with the correct novel object when speaker and child were jointly focused on one and the same object at the time of labeling. However, when faced with discrepant labeling (a speaker looked at and labeled a different object than the one on which children were focused), children with autism were prone to error—frequently, they mistakenly linked the new word with the object of their own focus, rather than with the object the speaker was actually intending to label. In contrast, normal children and non-autistic mentally handicapped children typically avoided such errors, tending to

© Blackwell Publishers Ltd. 2001

appropriately link the new word with the object toward which the speaker displayed social clues.

These findings hint that social deficits experienced by children with autism directly undercut their word learning. Given that the social skills involved are thought to be central to other aspects of knowledge acquisition (e.g., Baldwin, 2000; Tomasello, 1999), we might then expect individuals with autism to suffer difficulties in many arenas that affect many aspects of socio-cultural participation. But before we rush to embrace this relatively pessimistic view, we should consider the strength of the available evidence.

First, there is reason to question our initial construal of the Baron-Cohen et al. evidence. While children with autism in that study displayed higher levels of word learning errors in the context of discrepant labeling, it is not certain that these errors were due specifically to deficits in social understanding (Ganea, Harris, & Lillard, in preparation). Quite possibly, the errors arose because individuals with autism suffer from other information-processing difficulties. For example, discrepant labeling presents children with a processing-intensive learning task. They must note the discrepancy between their own and the speaker's focus of attention, disengage their attentional focus from the object of current interest, process the speaker's social clues, follow those clues, appreciate the significance of those clues for interpreting the new word, all the while processing phonological characteristics of the novel word as well as physical characteristics of the target object, and linking these representations in memory. This is a tall order. Perhaps children with autism can't cope with the processing demands, and for this reason are subject to a higher rate of confusions in the phase of linking representations of word and object in memory. Or perhaps children with autism understand the value of social clues for guiding word learning, but simply can't voluntarily disengage attention from the object of current interest, rendering them prone to word learning errors. If this were correct, we would want to interpret their difficulties in terms of deficits in inhibitory control, rather than deficits in social understanding, per se. And indeed, individuals with autism are known to have profound executive deficits as well as social deficits (Russell, 1997). Ganea and colleagues are currently investigating the viability of the deficits-in-executive-control account for the difficulties individuals with autism have with the discrepant labeling scenario.

The fact that non-autistic mentally handicapped children matched for chronological and mental age with the autism group were not prone to mapping errors in response to discrepant labeling seems to speak against the information-processing account. However, the information-processing deficits of non-autistic mentally handicapped children, while real, might differ from those of children with autism, rendering their good performance not definitive as a basis for comparison with the autistic group.

Part of the difficulty in interpreting the Baron-Cohen et al. findings as straightforward evidence that deficits in social understanding undercut word-learning is that we do not know some important things about the response to discrepant labeling that children with autism actually showed in that study. In particular, it is not clear from the study whether children with autism a) failed to gather relevant social clues (i.e., did they check the speaker's face on hearing a novel label at rates similar to those of normal and non-autistic mentally handicapped children?), or b) gathered such information, but failed to use it to effectively. In other words, we can't yet determine whether children with autism are disinclined to monitor language-relevant social clues, such as gaze direction, or whether they are simply at a loss about how to put these clues to work for

© Blackwell Publishers Ltd. 2001

the purpose of word learning. The research underway by Ganea et al. (in preparation) is designed to provide information on these important points, as well.

It is even possible that word learning errors made by children with autism stem from deficits in social *interest*, rather than deficits in social understanding. This is what we might expect if Hobson's (1989) and Sigman and Capps' (1997) account of autism is borne out. That is, individuals with autism may be disinclined to look at others or actively seek social information from others simply because they are not particularly affiliative. This could leave them prone to errors when discrepant labeling occurs, while such errors would be non-problematic if the onus were not on them to gather the relevant social information.

At present, then, we have no strong positive evidence indicating that skills for interpreting social clues are crucial for normal word-learning. In addition, the existence of other clinical populations who show deficits in social understanding but seemingly intact word learning would seem to pose a real challenge to the hypothesis. In particular, high-functioning individuals with autism, those with Asperger's syndrome, and those with right-hemisphere and non-verbal learning disabilities all show difficulties in processing social information (Harris, 1995; Yirmiya & Sigman, 1991). Yet difficulties in vocabulary acquisition—at least with regard to learning to *comprehend* new words—are not typically reported for these individuals and sometimes are explicitly denied (e.g., Rourke & Tsatsanis, 1996). For example, APA's DSM-IV diagnostic criteria for Asperger's disorder specifically states 'there is no clinically significant general delay in language (e.g., single words used by age 2 years, communicative phrases used by age 3 years).' (APA, 1994). It seems that word learning in these individuals remains relatively unaffected by their social understanding deficits. To be clear, it is not the case that this group has no language difficulties; our point is specifically that difficulties with *vocabulary acquisition* are not typically reported.

This said, it is important to recognize that, as yet, careful probes of word learning in individuals suffering these disorders have not been carried out. It is certainly conceivable that they are delayed in vocabulary acquisition, yet not so drastically delayed that this stands out to parents who are reporting retrospectively on their children's early development. Also conceivable is that these individuals experience a higher incidence of word-learning errors than children with intact social understanding, but that they are able to maintain normal levels of progress due to several other factors. These factors could include a) compensatory reliance on other sources of information about word-to-world links, and/or b) compensatory interaction strategies on the part of their caregivers, who, noting children's difficulties (perhaps even unconsciously), work hard to minimize occasion for error, by avoiding discrepant labeling, for instance.

Also, social understanding probably is not a monolithic skill, but rather a collection of component skills. These skills might be affected to some degree independently. Those who lack many of these skills—for instance, children with 'Kanner-type' (Frith, 1989) autism—may show profound difficulties in word learning, while others—such as high-functioning autistics and those with Asperger's or non-verbal learning disability—may retain some component skills, enabling them to avoid profound word-learning deficits.

In support of this latter possibility, individuals with Asperger's syndrome and those with autism who are high functioning typically perform well on standard measures of 'theory of mind'—questions probing a basic understanding of others' intentions, beliefs, knowledge, and the like (Harris, 1995; Baron-Cohen, Tager-Flusberg, & Cohen, 2000). Thus, they seem to possess an intact appreciation for the social phenomena we

© Blackwell Publishers Ltd. 2001 

have suggested are crucial for word learning. Their social deficits, then, while real, might not impact word learning to any significant degree. Developing new tasks with an eye to distinguishing possible different components of social skill and exploring their relevance to word learning will be necessary to probe these issues further.

Another potentially powerful, but as yet relatively untapped, source of information about relations between social understanding and word learning is individual variation within *typically-developing* children. Infants and young children within the typical range display marked differences in social affiliativeness, attentiveness to social signals, openness to regulation by others, executive control, and the like. If our proposal about an important link between social understanding and word learning is on the right track, we may see evidence for this link as a result of typical developmental variation. Along these lines, Brand (in preparation) has recently begun investigating whether the development of inhibitory control plays a role in enabling infants to cope with discrepant labeling. To achieve word-learning success when adults label something other than what infants are focused on, infants must inhibit their interest in their own object in favor of consulting the speaker's social clues and discovering the speaker's intended referent. There is reason to believe that normal infants achieve such inhibitory control on differing schedules, which may affect their progress in vocabulary acquisition.

Overall, then, existing evidence from children with autism is consistent with the view that social understanding is central to word learning, but there is much to be learned before a compelling case can be made. Capitalizing on individual variation within typical populations will also be important in rounding out the story.

## Processing Intentions On-line: Too Demanding to be of General Value?

We have made much of the scope and flexibility of normal infants' skill at inferring others' intentions. But is this something they can afford to engage in day in and day out, in each and every situation involving language that is new to them? To draw such inferences about intentions with any accuracy, infants must be acute observers of others' intonation, facial expressions and bodily motions, process these motions as they change over time in relation to objects and locations in the physical context, and register how the language produced corresponds with this whole dynamic flow of multi-modal information. This seems monumentally demanding in terms of processing resources, which we have reason to suspect infants might have in relatively short supply (e.g., Case, 1998).

Still, we should not write-off the possibility that this is just what infants do. Two lines of evidence would be helpful. First, we need to know more about the intentions adults reliably enact in the context of their language input to children. Perhaps they enact, or more likely foreground, a relatively small set of intentions again and again (intentions to eat, sleep, praise, engage in joint attention, warn, soothe, play peekaboo, etc.). If so, over time infants would be saved much inferential labor. Having come to interpret the small, high-lighted, set of adult intentions, and recognized their relative stability, infants would be faced with a processing task of reduced proportions. Others have suggested just such a scenario as important in language learning (e.g., Bruner, 1977, 1983; Fernald, 1989; Nelson, 1985), and recent research provides supportive evidence. For example, Pan, Imbens-Bailey, Winner, and Snow (1996) found that a small range of intentions, such as discussing a joint focus, directing the child's attention, and negotiating immediate activity, accounted for a large proportion (close

© Blackwell Publishers Ltd. 2001

to 80%) of parents' communicative intentions in parent-child interaction. And other research by Callanan and colleagues provides evidence that adults supply clear markers when their communicative intentions diverge from what we might call the 'default' (Callanan, 1985, 1989; Callanan & Sabbagh, under revision). Thus clues are available to help infants a) discern the intentions motivating parents' routine actions, b) detect when intention processing beyond the routine is needed, and b) discover what the substance of those divergent intentions might be.

Second, we need to know more about the nature of the system that supports infants' intentional inferences. The processing demands of inferring others' intentions depend in large part on the characteristics of the system that executes such inferences. If the system design readily supports the processing of intentions, then the processing demands would be reduced. Along these lines, it is at least plausible that infants come specially equipped with well-tuned skills for getting started in this arena. Attesting to this, within the first two months infants are already remarkably socially responsive—they smile when others exhibit a range of benign motions, expressions, and vocal intonations, and respond negatively to signs of harshness or even just disengagement from others (e.g., Toda & Fogel, 1993). It is not inconceivable that they already are recognizing, and distinguishing, at least some distinct intentions on others' part.

Also, our brains may be organized in ways that enable us to carry out intention detection with minor outlay of resources. In particular, several researchers (e.g., Driver, Davis, Ricciardelli, Kidd, Maxwell, & Baron-Cohen, 1999; Langton & Bruce, 1999, 2000) have recently presented evidence suggesting that, for adults at least, processing of gaze direction and pointing gestures is automatic and obligatory.

As yet, little neuroscience research has actually targeted these issues, but some intriguing findings have already emerged regarding motion-processing skills that may play a central role in human intention detection (Frith & Frith, 1999). Research by Perrett and colleagues (e.g., Perrett & Emery, 1994) reveals that certain brain regions in macaque monkeys respond selectively when the individual views hands interacting with objects, as when a hand reaches to grasp an object. These 'manipulation-tuned' cells respond only when the object is present and within contactable distance; that is, the identical pattern of motion yields no response if the targeted object is absent or unattainable. Moreover, the cells readily respond regardless of viewing orientation; for example, grasping can be witnessed from above, or in profile. Temporal cortex in the macaque monkey also contains cells that respond selectively to gaze, head, and/or body direction. These findings from macaques point to the brain being selectively tuned to processing just those aspects of bodily information valuable in discerning intentions. Adult macaque brains also appear to possess 'mirror neurons'—neurons that fire in response to actions, such as a hand moving to grasp an object, emitted by *either* the self or another (e.g., Goldman, in press; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Much remains to be learned about the precise function of all these cell types—manipulation-tuned cells, attention-direction cells, and mirror neurons—in the higher-order processing of action information, and about their developmental history. That is, mirror neurons may be functional as such from early in infancy, or alternatively, their characteristic response pattern may only emerge over the course of extended development. Regardless, their very existence in adulthood indicates that even macaque brains achieve an organization that would facilitate processing of action basic to the discernment of intentions. Of course, however, we have much to learn about the existence of analogous cells in the human brain.

© Blackwell Publishers Ltd. 2001

We also have much to learn about the mechanisms that enable infants to become skilled at discerning intentions in dynamic human action at such an early age. Intention detection seems to depend on a generative knowledge system that enables us to discern intentions even when people's actions are novel and are enacted in novel settings (e.g., Baldwin & Baird, in press). In question, then, is how infants so readily acquire this generative knowledge system. One approach to investigating this question is to examine infants' action processing skills directly. When people execute intentional actions they typically move rapidly through space, contact and manipulate numerous objects, moving their appendages and sensors in complex ways. The flow of motion is both dynamic and continuous, without obvious pauses to demarcate one action from the next. By 10–11 months infants possess some skill for parsing such continuous streams of intentional action in terms of relevant action units—units bounded by the initiation and completion of intentions (Baldwin, Baird, Saylor, & Clark, in press). These findings, along with others from researchers such as Gergely and colleagues (1995), Wellman and Phillips (in press), and Woodward and Somerville (in press) make a promising start at illuminating the genesis of social understanding.

## Conclusion

We began with 'no legal precedent,' an anecdote illustrating social understanding at play in aiding word learning. Current evidence is very strong that social understanding buoys word learning from a remarkably early age. Yet there is much we don't yet know. Do infants consult social clues because they are one among a useful collection of tools for word learning, or do they treat social clues as the pivot around which inferences about word meaning necessarily turn? Can those who suffer difficulty processing social information nevertheless progress normally in word learning, by virtue of compensatory strategies, or compensatory assistance from adults? What is the nature of the system that supports our facility at discerning the intentions motivating others' actions? We are delighted to find the field poised to identify answers to these important questions.

## References

Akhtar, N. & Tomasello, M. (1996). Twenty-four-month-old children learn words for absent objects and actions. *British Journal of Developmental Psychology*, *14*, 79–93.

Akhtar, N., Jipson, J., & Callanan, M. (in press). Learning words through overhearing. *Child Development*.

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, *67*, 635–645.

American Psychiatric Association, Committee on Nomenclature and Statistics. (1994). *Diagnostic and statistical manual of mental disorders*, 4th ed., Washington, DC.

Baldwin, D. A., Bill, B., & Ontai, L. L. (May, 1996). *Infants' tendency to monitor others' gaze: Is it rooted in intentional understanding or a result of simple orienting?* Paper presented at the International Conference on Infant Studies, Providence, Rhode Island.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*, 875–890.

Baldwin, D. A. (1993a). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, *20*, 395–418.

Baldwin, D. A. (1993b). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*, 832–843.

Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. *Current Directions in Psychological Science*, *9*, 40–45.

© Blackwell Publishers Ltd. 2001 

Baldwin, D. A. & Baird, J. A. (in press). Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*.

Baldwin, D. A., Baird, J. A., Saylor, M., & Clark, M. A. (in press). Infants parse dynamic human action. *Child Development*.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, *67*, 3135–3153.

Baldwin, D. A. & Moses, L. J. (1996). The ontogeny of social information-gathering. *Child Development*, *67*, 1915–1939.

Baldwin, D. A. & Tomasello, M. (1998). Word learning: A window on early pragmatic understanding. In E. Clark (Ed.), *Proceedings of the Twenty-ninth Annual Child Language Research Forum*, *29*, (pp. 3–23). Cambridge, UK: Cambridge University Press.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze (SDG) strategy to crack the code of language? *Child Development*, *68*, 48–57.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.). (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford: Oxford University Press.

Bloom, P. (2000). How children learn the meanings of words. Cambridge, MA: MIT Press.

Brand, R. (in preparation). Relations between inhibitory control and early word learning. Unpublished doctoral dissertation. University of Oregon, Eugene, OR.

Bruner, J. (1977). Early social interaction and language acquisition. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 271–289). New York: Academic Press.

Bruner, J. (1983). *Child's talk: Learning to use language*. New York: W. W. Norton.

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development*, *56*, 508–523.

Callanan, M. A. (1989). Development of object categories and inclusion relations: Preschooler's hypotheses about word meaning. *Developmental Psychology*, *25*, 207–216.

Callanan, M. & Sabbagh, M. (under revision). *Multiple labels for objects in conversations between young children and their parents: Pragmatic basis of a word-learning principle*. University of California at Santa Cruz.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- to 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*, 315–330.

Case, R. (1998). The development of conceptual structures. In W. Damon (Series Ed.) & D. Kuhn, & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception and language* (5th ed.) (pp. 745–800). New York: Wiley.

Diesendruck, G., Markson, L., Akhtar, N., & Reuder, A. (2001). 'Two-year-olds' sensitivity to speakers' intention: an alternative account of Samuelson and Smith.' Manuscript under review.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*, 509–540.

Dunham, P. J., Dunham, F., & Curwin, A. (1993). Joint attentional states and lexical acquisition at 18 months. *Developmental Psychology*, *29*, 827–831.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, *60*, 1497–1510.

Frith, U. (1989). *Autism: Explaining the enigma*, Oxford, U.K.: Basil Blackwell.

Frith, C. D. & Frith, U. (1999). Interacting minds—a biological basis. *Science*, *286*, 1692–1695.

Ganea, P., Harris, P., & Lillard, A. (in preparation). *Joint attention and word learning in children with autism.* University of Virginia, Charlottesville, VA.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.

Goldman, A. I. (in press). Desire, intention, and the simulation theory. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*. Cambridge, MA: MIT Press.

Harris, J. C. (1995). *Developmental neuropsychiatry, Volume I: The fundamentals; Volume II: Assessment, diagnosis, and treatment of developmental disorders*. New York, NY: Oxford University Press.

© Blackwell Publishers Ltd. 2001

Hobson, R. P. (1989). Beyond cognition: a theory of autism. In G. Dawson (Ed.), *Autism: Nature, diagnosis, and treatment*, (pp. 22–48). New York: Guildford Press.

Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65* (3).

Langton, S. R. H. & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*, 541–567.

Langton, S. R. H. & Bruce, V. (2000). You must see the point: Automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 747–757.

Meltzoff, A. & Brooks, P. (in press). In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition. Cambridge, MA:* MIT Press.

Meltzoff, A. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 1–16.

Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental Psychology*, *35*, 60–68.

Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (in press). Evidence for referential understanding in the emotions domain at 12 and 18 months. *Child Development*.

Nelson, K. (1985). *Making sense: The acquisition of shared meaning*. New York: Academic Press.

Pan, B. A., Imbens-Bailey, A., Winner, K., & Snow, C. (1996). Communicative intents expressed by parents in interaction with young children. *Merrill-Palmer Quarterly*. *42*, 248–266.

Perrett, D. I. & Emery, N. J. (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. *Current Psychology of Cognition*, *13*, 683–694.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rourke, B. P. & Tsatsanis, K. D. (1996). Syndrome of nonverbal learning disabilities: Psycholinguistic assets and deficits. *Topics in Language Disorders*, *16*, 30–44.

Russell, J. (Ed.). (1997). *Autism as an executive disorder*. Oxford: Oxford University Press.

Sabbagh, M. A. & Baldwin, D. A. (in press). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*.

Samuelson, L. K. & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, *69*, 94–104.

Saylor, M., Sabbagh, M. A., & Baldwin, D. A. (under revision). Implicit contrast is a pragmatic cue to word meaning. *Child Development*.

Schafer, G. & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, *69*, 309–320.

Sigman, M. & Capps, L. (1997). *Children with autism: A developmental perspective*. Cambridge, MA: Harvard University Press.

Stager, C. L. & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382.

Toda, S. & Fogel, A. (1993). Infant response to the still-face situation at 3 and 6 months. *Developmental Psychology*, *29*, 532–538.

Tomasello, M. (1992). The social bases of language development. *Social Development*, *1*, 67–87.

Tomasello, M. (1999). *Cultural Origins of Human Cognition*. Harvard: Harvard University Press.

Tomasello, M. & Barton, M. (1994). Learning words in non-ostensive contexts. *Developmental Psychology*, *30*, 639–650.

Tomasello, M. & Farrar, J. (1986). Joint attention and early language. *Child Development*, *56*, 1454–1463.

Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language*, *23*, 157–176.

Yirmiya, N. & Sigman, M. D. (1991). High-functioning individuals with autism: Diagnosis, empirical findings, and theoretical issues. *Clinical Psychology Review*, *11*, 669–683.

Wellman, H. M. & Phillips, A. T. (in press). Developing intentional understandings. In B. Malle, et al. (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*, MIT Press.

© Blackwell Publishers Ltd. 2001

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, *34*, 1289–1309.

Woodward, A. L. & Somerville, J. A. (in press). Twelve-month-olds interpret action in context. *Psychological Science.*

## Acknowledgments

Preparation of this article was supported by a National Science Foundation NYI Award to the first author. Our sincere thanks to Mark Sabbagh and Megan Saylor for thought-provoking discussion, and to Alison Imbens-Bailey and an anonymous reviewer for helpful comments on earlier drafts of this article.

© Blackwell Publishers Ltd. 2001