Abstract and Keywords

This chapter explores the early stages of pragmatic language acquisition before taking up the issue of syntax and semantics. It suggests that the relation between language and theory of mind is different depending on which aspect of social understanding is at issue. In particular, it argues that an appreciation of other persons as intentional agents—the first level of social understanding—is a prerequisite for language acquisition. The chapter presents evidence from a training study aimed at developing false-belief reasoning to show that both conversation about deceptive objects and training on the syntax of complementation (in the absence of deceptive objects) promote three-year-olds' false-belief understanding. The largest training effect occurred in a condition that combined conversation and complements. The fact that
manipulating the deceptive objects without any conversation about them was ineffective leads to the conclusion that language is a necessary condition for children to make progress in their understanding of false beliefs, lending support to the claim that language plays a causal role in the ontogeny of social understanding.

But there is nothing odd about the product of a given process contributing to, or even becoming an essential factor in, the further development of that process.

—George Herbert Mead, 1934, p. 226

“Language” and “theory of mind” are both global terms that cover a wide array of phenomena. Therefore, to be precise about a link between the two, it is first necessary to clarify which aspects of language and which aspects of theory of mind are at issue. We focus here less on Language as an abstract representational medium and more on linguistic communication as the process by which people use linguistic conventions to pursue communicative goals. And we focus not on Theory of Mind, narrowly defined as the understanding of false beliefs, but rather on children’s growing understanding of other persons more generally—what we shall call simply social understanding.

In particular, we are concerned here with two levels of social understanding that we have previously differentiated (Tomasello, 1999): (1) the understanding of other persons as intentional agents, whose behavior is governed by goals and perceptions, and (2) the understanding of other persons as mental agents, whose behavior is governed by goals (desires) and beliefs, including ones that are false. Our central claim is that the relationship between linguistic communication and social understanding is different depending on which of these two levels is involved. Fundamentally, understanding others as intentional agents is necessary for human beings to comprehend and acquire the use of linguistic symbols and
conventions in the first place. So, in this case, language arises from certain skills of social understanding—and this would be the case both phylogenetically (historically) and ontogenetically. But then the use of these intentionally constituted linguistic conventions in communicative interactions (discourse) with other persons leads young children (and perhaps led prehistoric humans) to an understanding of beliefs. More specifically, children come to understand beliefs as a result of participating in various kinds of linguistic interactions in which their interlocutor shifts perspectives in various ways—including taking various perspectives and modal attitudes both on the child’s and on her own linguistically expressed beliefs (e.g., in prepositional attitude constructions). In the case of understanding beliefs, then, linguistic interactions help to create new forms of social understanding.

Early Intention Reading and Language

Dogs understand verbal commands, parrots produce English words and phrases, and vervet monkeys give different alarm calls for different predators. Why are these not language? They are not language because they do not have the requisite psychological dimension. These animals are not trying to understand or manipulate the intentional or mental states of others; rather, they are focused on the behavior (or perhaps motivational states) of others. They are therefore comprehending and using communicative signals—either not learned or learned associatively—not linguistic symbols. Operating with linguistic symbols requires a certain kind of social understanding.

We can see this most clearly if we look at the process of word learning. Most theories of word learning hold that a linguistic symbol is simply a sound (or possibly a hand sign) that “stands for” something in the world (or else a concept). What it means for one thing to stand for another is never really addressed. The implicit assumption is that it is just an association—which makes it pretty much like what dogs and parrots do. But if we look at children’s earliest comprehension and production of real-life linguistic utterances, we see that something very different is going on. The child encounters an adult making
funny noises at her. What is she to make of this odd behavior? She could simply associate these noises with ongoing experiences, like a dog. But if she understands the other person as an intentional being, she will attempt to determine the purpose for which that person is making these funny noises. One possibility is that the adult is attempting to communicate with her, and so the child must attempt to determine her communicative intentions.

This is not a process of association. This is a process of establishing joint attention, and it relies on intention reading—and indeed a special form of intention reading at that (Tomasello, 1999, 2001). It requires the child to understand not just the adult’s intentions to some outside entity but also his intentions toward her (the child’s) attention to some outside entity—that is, his communicative intentions (Tomasello, 1998). This is what a linguistic convention or symbol is. It is a sound (or other behavior) that two or more individuals use with each other to direct each other’s attention and thereby to share attention—and they both know this is what they are doing. If you do not have this, then you may have something like the family pet knowing that the sound “dinner” means that food is coming, but you do not have an intersubjectively understood linguistic symbol used to follow into, direct, and share attention with other persons. This is the essence of what has been called the social-pragmatic theory of language (especially word) learning (Bruner, 1983; Nelson, 1985; Tomasello 1992).

This way of viewing early language acquisition explains in a way that no other theory of early language does—especially all of those relying on simple association—why language acquisition begins when it does in the months just after the child’s first birthday. Thus, children begin to develop nonverbal joint attentional skills at around 9 to 12 months of age, including such things as following the gaze direction and gestures of adults, imitating adult actions on objects, and directing adult attention to outside objects using various kinds of gestural signals (see Tomasello, 1995, for a review). Many children also show their first signs of comprehending language at this same age, with the first linguistic productions coming
soon after (Fenson, Dale, Reznick, et al., 1994). Most important, in a recent longitudinal study, Carpenter, Nagell, and Tomasello (1998) found that children’s comprehension and production of language correlated highly with their skills of joint attentional engagement with their mothers. The reason that linguistic skills are so highly correlated with joint attentional skills is that language is nothing more than another type—albeit a very special type—of joint attentional skill; people use language to influence and manipulate one another’s attention.

The social understanding required for word learning can be seen especially clearly in some recent studies in which children learn new words in the ongoing flow of social interaction in which both they and the adult are trying to do things. For example, here are five different situations—all requiring some kind of intention reading—in which 18- to 24-month-old children learned new words with some facility.

- In the context of a finding game, an adult announced her intention to “find the toma” and then searched in a row of buckets, all of which contained novel objects. Sometimes she found it in the first bucket searched. Sometimes, however, she had to search longer, rejecting unwanted objects by scowling at them and replacing them in their buckets until she found the one she wanted. Children of 18 and 24 months of age learned the new word for the object the adult intended (p.248) to find (indicated by a smile and termination of search) regardless of whether or how many objects were rejected during the search process (Tomasello & Barton, 1994; Tomasello, Strosberg, & Akhtar, 1996).

- Also in the context of a finding game, an adult had the child find four different objects in four different hiding places, one of which was a very distinctive toy barn. Once the child had learned which objects went with which places, the adult announced her intention to “find the gazzzer.” She then went to the toy barn, but it turned out to be “locked.” She thus frowned at the barn and then proceeded to another hiding place, saying, “Let’s see what else we can find” (taking out an object with a smile). Later, children of 18 and 24 months of age demonstrated that they had
learned “gazzer” for the object they knew the experimenter wanted in the barn, even though they had not seen the object after they heard the new word and even though the adult had frowned at the barn and smiled at a distractor object (Akhtar & Tomasello, 1996; Tomasello et al., 1996).

• An adult set up a script with the child in which a novel action was performed always and only with a particular toy character (e.g., Big Bird on a swing, with other character-action pairings demonstrated, as well). She then picked up Big Bird and announced, “Let’s meek Big Bird,” but the swing was nowhere to be found—so the action was not performed. Later, using a different character, 24-month-old children demonstrated their understanding of the new verb even though they had never seen the referent action performed after the novel verb was introduced (Akhtar & Tomasello, 1996).

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

• A child, her mother, and an adult played together with three novel objects. The mother then left the room. A fourth object was brought out, and the child and adult played with it, noting the mother’s absence. When the mother returned to the room, she looked at the four objects together and exclaimed, “Oh look! A modi! A modi!” Understanding that the mother would not be excited about the objects she had already played with previously but that she very well might be excited about the object she was seeing for the first time, 24-month-old children learned the new word for the object the mother had not seen previously (Akhtar, Carpenter, & Tomasello, 1996).

No theory of associative learning can explain these results. Although any one of these studies might be explained in other ways, when they are considered as a group the most plausible explanation is that, by the time they are 18 to 24 months of age, children have developed a deep and flexible
understanding of other persons as intentional beings, and so they are quite skillful at determining the adult’s communicative intentions in a wide variety of relatively novel communicative situations.

And so linguistic communication is made possible by a certain form of social understanding, namely the understanding of the intentions and communicative intentions of other persons. Other animal species do not operate with this kind of understanding, and so while they can learn to associate sounds with experiences, they do not acquire the use of linguistic symbols as social conventions for manipulating the intentional and mental states of others. So, at this level of analysis, social understanding enables, indeed is in an important sense constitutive of, language.

Language and the Development of False-Belief Understanding

Once children have begun communicating with other people using historically evolved linguistic conventions, they are exposed to all of the different categories and perspectives that previous individuals in their linguistic community have found it useful to employ communicatively. And, in Vygotskian fashion, as they internalize these linguistic symbols, they begin to develop concepts and social-cognitive skills similar to those of other persons in their linguistic community, as well.

A social-cognitive skill currently of great interest to many developmentalists, including many of the authors in this volume, is the understanding of false beliefs. Arguably, the acquisition of language plays a central role in the ontogeny of this understanding. Thus, language development and false-belief understanding are relatively strongly correlated, as established by Dunn, Brown, Slomkowski, Tesla, and Youngblade (1991), Astington and Jenkins (1999), Gale, de Villiers, de Villiers, and Pyers (1996), de Villiers and de Villiers (2000), Watson, Painter, and Bornstein (2002), and Farrar and Maag (2002).

In addition, some investigators have used training studies in an attempt to go beyond correlations and to demonstrate specific causal relations between children’s experience during
training and their false-belief understanding. Relevant studies of this kind are Appleton and Reddy (1996), Swettenham (1996), Slaughter and Gopnik (1996), Slaughter (1998), McGregor, Whiten, and Blackburn (1998), Clements, Rustin, and McCallum (2000), and Hale and Tager-Flusberg (2003). The type of training was different in each of these studies, but in all cases in the key conditions children experienced some kind of deceptive scenario involving issues of appearance-reality and/or false belief (including in some cases training on false-belief tasks directly), along with linguistic descriptions of that scenario, typically including mental-state talk. The problem is that none of these studies had a control condition in which children experienced some kind of deceptive scenario during training but without any linguistic description at all. Such a control condition is necessary to unconfound deceptive experience and language and thus to determine whether language influences false-belief understanding over and above training involving deceptive experiences.

Beyond the question of whether or not language plays a role in false-belief understanding, very little research has been aimed at identifying specifically the nature of this role. There are four global hypotheses. The first is that language has no special role to play. The idea is that children are constantly forming theories about other people and their minds and that any and all data are relevant. Data from linguistic sources may be used, but they have no special status. Although it is unclear whether anyone espouses this view in its pure form, the theory-theory certainly tends in this direction (Gopnik & Wellman, 1992). The second hypothesis is that learning mental-state terms such as think, know, and believe plays a key role in the development of false-belief understanding (see, e.g., Olson, 1988). The idea here is that these particular linguistic symbols are used by adults to indicate the relevant mental states, and so in learning the referents of these terms children form, in Whorfian fashion, the relevant concepts. Again, it is unclear whether anyone espouses this as the exclusive, or even as the single most important factor involved. But Bartsch and Wellman (1995) discuss in detail the
possible importance of this language learning process, and Astington (2000) also seems to accord it some importance.

Third, de Villiers and de Villiers (2000; see also Gale et al., 1996) proposed that the syntax, that is, the grammatical form, of the way adults talk about beliefs and related mental states provides children with a necessary representational format for dealing with false beliefs. Specifically, what is said to be crucial is the syntax of complementation, in which a sentence takes a full clause as its object complement (sentential complements). For example, consider Peter thinks Mommy's home or You know that I'm not coming to the party. This hypothesis has recently received support from a training study in which children trained with sentential complement sentences subsequently improved in their false-belief understanding (Hale & Tager-Flusberg, 2003). However, it is important to point out that the sentences in this training study were given in talking about deceptive experiences, and so it is possible that it was these experiences, and not the sentences themselves, that led to the increase in false-belief understanding (i.e., there was no control condition without deceptive experience).

A fourth hypothesis, put forward by Harris (1996, 1999), proposed that it is not the semantic content of mental-state terms or the syntax of complementation that are key linguistic influences on the understanding of false beliefs but rather that the key is the process of linguistic interchange that children experience in discourse with other people. The idea here is that the whole notion of “belief” as a mental state makes sense only in the context of alternative possible beliefs about a situation, including one that is “true” (implying that others may be false). Harris claims that it is in the to-and-fro of discourse that the child comes to appreciate that other people know things she does not know, that they do not know things they ought to know, and that they have different perspectives on things (see also Tomasello, 1999; Siegal, 1999). Most training studies have employed rich discourse interactions as a part of the training (see preceding discussion), and so it is possible that these studies provide support for the discourse view. But, again, these training
conditions typically involved a number of other factors in addition, including mental-state terms and sentential complements.

We report here two training studies. The first study attempted to establish both (1) whether language causally influences false-belief understanding (with the appropriate control group), and (2) if so, the nature of this influence with respect to the various hypotheses previously proposed (with the appropriate comparison groups).

Study v Training False-Belief Understanding
In this study, 3-year-old German children (mean age = 42.6 months) experienced pretests, intervention training, and then posttests. Children who had not yet acquired false-belief understanding (and who showed linguistic development within the norms of their age group) took part in the intervention training. All of the details are reported in Lohmann and Tomasello (2003), from which the following section is abstracted.

Pretests
Each child was given three pretests: (a) a vocabulary test, (b) a false-belief task (the representational change task; Gopnik & Astington, 1988), and (c) two tests of the child’s comprehension of sentential complement constructions (a modification of Swettenham’s (1996) Tom test, and the memory for complements test of de Villiers and de Villiers, 2000; see also J. de Villiers, this volume, chapter 10).

Training
There were four different training conditions. All involved adult-child interactions with deceptive objects (e.g., children see an object that looks like an apple but is really a candle). Training in the different groups was as follows:

(p.252) Full Training
In this condition, the deceptive aspect of the training objects was highlighted and the experimenter (E) talked about this using either mental-state verbs (think, know) or communication verbs (say) within sentential complement constructions. Thus, for each object, E showed the child the
object and asked her what she thought the object was, using psychological verbs (e.g., *think*) and sentential complement constructions. Then the object was handed to the child so that she could see its “real” function, which was highlighted by E. Then the child was asked to recall her previous belief and her current knowledge of the object’s nature, which was possibly corrected and summarized once more by E. The child was also asked to predict a third person’s reaction to the deceptive object. A hand puppet, Schnuffi, was used to represent a third person, and the child watched as E informed Schnuffi of the real function of the object. Children could observe the surprise reactions of the puppet and assist the puppet in finding out the real function of the object. Finally, the children were asked about the puppet’s new (changed) belief about the object. (To see whether specific verbs made a difference, half the children in this condition were trained using only mental-state verbs such as *think* and *know*, and half were trained using only the communication verb *say*. Since these options turned out to produce the same results, they were subsequently collapsed).

**Discourse Only Training**

In this condition, the deceptive aspect of the training objects was highlighted, but E did this without using either mental-state verbs or sentential complement constructions. Thus, for example, instead of asking, “What do you think this is?” E asked the child, “What is this?”; instead of saying, “A flower,” E said, “A flower.”

**No Language Training**

In this condition, the deceptive aspect of the training objects was highlighted, but E did this essentially non verbally. Thus, children were first shown an object, and E said, “Look!” and then their attention was drawn to the real function by E’s showing it and saying, “But now look!” The appearance/reality distinction of the objects was highlighted twice by E with appropriate nonverbal emotional expressions; no questions were asked, and no feedback was given to the child. For the third-person perspective, the hand puppet was brought out and shown the object: “Look, Schnuffi.” The child observed the puppet’s reactions to the object, which showed surprise
reactions to the real function of the object: “Oh!,” “All right,” and the like.

Sentential Complements Only Training

In this condition, the deceptive aspect of the training objects was not highlighted for the child in any way. E simply talked about them as normal objects, using mental verbs or communication verbs and sentential complements. Four short stories were designed so that E could talk about the objects without referring to their deceptive nature. Thus, children were asked what they thought the object was and about certain attributes of the object, while the deceptive nature of the objects was never revealed. To avoid contrasting mental states, children’s answers were not contradicted. In order to stimulate the children’s acquisition/comprehension of sentential embeddings, the children were also asked to help E clarify what the protagonist of the story had done or said. For example, a handpuppet Ernie says: “This chair belongs to my grandfather. I know that!” E then asks the child: “What does Ernie know?” To answer the question appropriately, children had to use one clause as a sentential complement of the other: “He knows that the chair belongs to his grandfather.” The same number of sentential complement sentences with mental verbs was used in feedback and questions as in the Full Training group, but without referring to any kind of contrasting deceptive experiences.

During training, each child interacted with an adult experimenter on each of four occasions within a two-week period. The basic training procedure was modeled on that of Slaughter and Gopnik (1996). Children in all training groups were exposed to 16 objects. Each was brought out singly and replaced after discussion of it was completed. Order was randomized across children (except in the Sentential Complements Training condition, in which a fixed order was used so as to tell a story). Twelve of the objects were deceptive objects in the sense that on first glance they appeared to be one thing (e.g., a flower) but on closer inspection they had another function (e.g., a writing pen). Four objects did not have a deceptive aspect and were used randomly among the 12 deceptive ones to prevent children
from expecting every object we showed them to be something other than what it appeared to be. Training for each group consisted of discussion of each object, with the experimenter providing feedback and corrections to the child’s comments where appropriate (except for the No Language Training group, where no feedback was given and no questions asked). The child’s first session consisted of the pretests and training with 3 deceptive objects. In the second session, 5 deceptive objects were discussed; in the third session, 4 deceptive objects were discussed; and in the final session the posttests were administered.

**Posttests**

As outcome, we measured three different types of false-belief understanding (one appearance-reality test similar to the training and two transfer tests).

- The appearance-reality task with deceptive objects (Flavell, 1986) was similar in format to the training procedure. In order to look for influences of the different training elements on children’s understanding of the appearance-reality distinction, the test was scored with three points: one point was given for the reality, one for the appearance, and one for the third-person prediction question.

(p.254)  
- The representational change task (Gopnik & Astington, 1988) was similar in structure to the pretest, but with different content (this served as one transfer task—one point was given for each of the two questions concerning the child’s own and a third person’s mental state).

- The change-of-location task (Wimmer & Perner, 1983) served as another test of transfer of training (one point was given for a correct answer to the test question “Where will the protagonist look for the object?”).

Children were also given sentential complement posttests. The format of the complement pretests was used again for the posttest, but with different story content in both cases.
Results
Preliminary analyses established that all four training groups were equivalent on all relevant measures at pretest. With regard to posttests, a significant effect of group was found for the first transfer test of false-belief understanding: Representational Change Task. The Full Training group outperformed each of the other groups on this task at posttest, and none of the paired comparisons among these other three groups revealed significant differences. Converted to percentages, the Full Training group was correct on average on 75% of the posttest questions; the Discourse Only and Sentential Complements Only groups averaged about 40% correct; and the No Language group averaged only about 25% correct answers. No difference between first- and third-person test questions was found. Pretest and posttest scores on the Representational Change Task are presented in table 12.1. Testing each group individually against chance, it was found that all groups except the No Language group increased their performance significantly compared to their pretest scores.

The other transfer task is the Location Change task. There was also a significant group difference for this task. Post hoc tests showed that the Full Training group was significantly better on the Location Change task than the No Language group, with no other groups differing from one another.

The Appearance-Reality task measured outcome on a task similar in structure to the training itself. No significant effect of group was found on the sum of questions in this “training-related” false-belief task. It appears that any training with experience of appearance-reality objects helped children to pass this test. But it should also be noted that even children with no deceptive experience but linguistic training on sentential complements did well on this task; this may be a result of the fact that the test uses test questions with sentential embeddings (similar to training). However, further analyses revealed group differences on the third-person prediction question of the test. Results of a logistic regression showed a significant group effect for this question. Post hoc tests showed that the Full Training led to significantly more correct answers on this question than the
No Language Training, once again duplicating the finding for the Representational Change task.

<table>
<thead>
<tr>
<th>Table 12.1 Means Scores on the Representational Change Task at Pretest and Posttest for Each Training Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Pretest False Belief (0–1) mean:</td>
</tr>
<tr>
<td>Posttest False Belief (0–2) mean:</td>
</tr>
</tbody>
</table>

Note: Children could not get 2 points in the pretest because passing the test was used as an exclusion criterion for the training.

* indicates a significant improvement from pretest to posttest score. Sign test:

(*) = p < .05

(**) = p < .01

These analyses of the individual posttests provided evidence that the Full Training condition was most effective in producing social-cognitive gains, but the Discourse Only and Sentential Complements Only also produced gains on some tasks. In a final, most general analysis, we summed across all false-belief posttest tasks to look for an overall effect of training. This analysis confirmed a group effect on false-belief understanding between the training groups. Figure 12.1 shows that the Full Training group performed better at posttest than each of the other groups on the sum of false-belief scores.
We also looked in more detail at the different training conditions, with a focus on (1) the Discourse Only group and how it compared with the others, and (2) the Sentential Complements Only group and how it compared with the others. First, the Discourse Only group provided children with differing perspectives on the deceptive objects using nouns, whereas the Full Training group did something similar, but using the language of sentential complements with mental verbs. The comparison between these groups is thus informative about the role of sentential complements over and above perspective-shifting discourse. Contrasts of the Full Training against the Discourse Only condition showed a significant effect of the Full Training. Thus, the use of sentential complements facilitated children’s false-belief understanding beyond that provided by the Discourse Only condition. Nevertheless, a planned contrast between the Discourse Only and the Sentential Complements groups found no difference, suggesting that the deceptive experience in the Full Training condition was an important factor, as well. In a final planned contrast, the Discourse Only group did not differ significantly from the No Language group.

Second, the Sentential Complements group was similar to the Full Training group, except that there was no deceptive experience involved. That is, the children in the Full Training group experienced deceptive objects and perspective-shifting talk about them (including both first- and third-person perspectives), whereas children in the Sentential Complements group did not experience deceptive objects or talk of contrasting mental states during the training. The results of a planned orthogonal contrast showed that the group with deceptive experience
(Full Training) outperformed the group without deceptive experience (Sentential Complements). Thus, the experience of changing perspectives on deceptive objects seems to be an important factor in the acquisition of false-belief understanding. Comparing the Sentential Complements with the No Language condition again revealed a significant difference in favor of the Sentential Complement training. Moreover, combining those conditions that used sentential complements in the training (Full Training and Sentential Complements) and comparing them with those conditions that did not (Discourse and No Language) also showed that the groups with sentential complements in the input outperformed the groups that did not use this specific linguistic construction in the training. The use of sentential complements thus seems to be an important factor in the acquisition of false-belief understanding independent of deceptive experience and perspective-shifting discourse.

Supportive of this last finding, another important result is that the use of sentential complements in training was correlated with improved scores on false-belief understanding. Analysis of children’s improvement on the sentential complement tasks at post-test showed that the most improvement came in the Sentential Complements training group. More importantly, table 12.2 shows the correlation between the change scores of false-belief understanding and the change scores on sentential complements (Test 1) of all training groups. A significant correlation was found for the Sentential Complements training group. This indicates that those children who improved their linguistic skills also improved their false-belief scores in this condition. In contrast, in the Full Training group, children’s improvement in false-belief understanding was not reliably associated with improved linguistic scores, presumably indicating the role of other effective factors (such as perspective-shifting discourse) in this condition.

Discussion
This study had three main findings. First, language was a necessary condition for young children to make progress in false-belief understanding. Simply experiencing deceptive objects was not sufficient; children needed to have that
experience structured by some language from other persons—for example, different nouns indicating different possible perspectives on these objects (mediated perspective-shifting discourse). Second, training in the syntax of sentential complements, including mental-state predicates as matrix verbs, was sufficient by itself to facilitate children’s false-belief understanding. This effect was evident even in a condition in which children had no experience with deceptive objects. Third, these two effects—of perspective-switching discourse and sentential complement syntax—seem to be relatively independent of each other. The strongest facilitator of children’s false-belief understanding in this study was a training condition that incorporated both of these factors, and the correlational findings provided further support for the independence of these two factors.

Table 12.2 Correlation of Pretest to Posttest Improvements on Both False-Belief and Sentential Complements Understanding for Each Training Group

<table>
<thead>
<tr>
<th></th>
<th>Change Scores on False Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Training</td>
</tr>
<tr>
<td>Change scores on sentential</td>
<td>r = -.18</td>
</tr>
<tr>
<td>complements</td>
<td>p = .22</td>
</tr>
</tbody>
</table>

The difference between the No Language training and the Discourse Only condition suggests that in our training it was the explicit labeling of the different perspectives that helped children to learn from the training. In the Discourse Only condition, the two perspectives were encoded in contentful linguistic symbols, such as “First it is a flower, and now it is a pen.” And so the effect of language had to do with the adult using conventionalized symbols (mainly in the form of common nouns) to highlight the different perspectives. It is important that the explicit encoding in this training condition did not involve reference to mental states themselves—nor did
It use any sentential complements—and so the effective factor in this condition was the process of perspective-shifting discourse, rather than any explicit reference to mental states using mental-state language or any syntactic format for symbolizing propositional attitudes.

Sentential complements (even without experience with deceptive objects or perspective-shifting discourse) were also an important factor. In the only other training study to investigate this question (Hale & Tager-Flusberg, 2003), the training of sentential complements always occurred in conjunction with deceptive experience, and so the current findings are the first to establish the important role of sentential complement syntax by itself in promoting the understanding of false beliefs. The current findings thus support the hypothesis of de Villiers and de Villiers (2000) that sentential complement constructions provide children with a convenient (if not necessary) representational format for conceptualizing and talking about false beliefs. It is also important to note that the sentential complement sentences in this training condition contained mental-state verbs, and so these might have played an important role in producing the training effect. However, in an explicit comparison of the two versions of the Full Training condition—one containing mental-state verbs (e.g., think, know) and one containing a communication verb (say)—no difference was found. This is at least indirect evidence that the main effect in the Sentential Complements condition was primarily tied not to the semantics of mental verbs but to the structure of sentential complement syntax (e.g., see Diessel & Tomasello, 2001). However, although a sufficient factor for false-belief understanding in this study, sentential complement syntax might not be a necessary factor for this understanding (see also Perner, Sprung, Zauner, & Haider, 2003; Perner, Zauner, & Sprung, this volume, chapter 11).

In virtually all analyses, the largest training effect was observed in the Full Training condition, which, like most previous training studies, contained both perspective-shifting discourse and sentential complement syntax, but no more overall talk than the other conditions. In combination with the findings that demonstrate the effectiveness of these two
factors by themselves, the superiority of the Full Training condition implies that both perspective-shifting discourse and sentential complement syntax make relatively independent contributions to children’s false-belief understanding.

The current study thus provides the strongest evidence to date that language plays a central role in children’s development of false-belief understanding. Specifically, perspective-switching discourse that uses contentful linguistic symbols (not necessarily mental-state language) and the ready availability of sentential complement syntax as a representational format both seem to make independently important contributions to the ontogenetic process. Further evidence for this general proposal comes from research with profoundly deaf children born to hearing families, who have almost no available means in their early years of conversing with their hearing family members, especially about topics such as mental states, which may have no obvious visual referent (de Villiers & deVilliers, 2000; Gale et al., 1996; Peterson & Siegal, 1995, 1998, 1999, 2000). These children, who are of normal intelligence, presumably experience as many situations as typically developing children do in which they observe others in surprise reactions or experience their own false beliefs. But these children struggle with false-belief tasks up to the age of 16 years. In contrast, deaf children born to signing parents, who share a communicative system and thus have much richer linguistic experiences, develop concepts of false belief at the same age as do hearing children (Peterson & Siegal, 1999, 2000).

Study 2: The Difference between First- and Third-Person Experience

Does the facilitating effect of perspective-switching discourse and sentential complement syntax hold also for younger children? In order to investigate this question, the following study used a modified version of the training procedure in Study 1 with children just turning 3 years of age. Another question is the possible role of self versus others’ experience—that is, whether children need themselves to engage in discourse about deceptive experiences or whether they can merely observe others in these kinds of interactions. On the basis of the claims of simulation theory (Harris, 1991), it might
be suggested that children would benefit most from having deceptive experience and talking about it themselves. Therefore, in this study, first- and third-person experience was separated into two conditions.

**Method**

Forty-eight German children between 2 years 10 months and 3 years 2 months of age (M = 3;0) participated in the study. All children failed a pretest of false-belief understanding before training began. As in Study 1, training consisted of a pretest, three training units with the first one starting immediately after pretest, and finally a posttest. Each child was given at pretest the same false-belief task and vocabulary test that was used in Study 1. In addition, all children received practice in forced-choice questions, which were regarded as central to the training. For this practice, children were shown 6 pictures of common content and then each time asked a forced-choice question with the correct answer (p.260) alternately mentioned at the beginning or at the end. For instance, a picture of a dog was presented, and then the child was asked: “Is this a dog or is this a cat?” After the training (described later), each child was given the same 3 false-belief posttests that were used in Study 1.

The training procedure was a modification of the training in Study 1. The structure was parallel in length and feedback to the most complex training condition of Study 1, the Full Training condition. As in Study 1, children were exposed to a total of 12 deceptive and 4 nondeceptive objects within the three training units. However, two training conditions were designed, a first-person and a third-person condition, as described in this section.

First the experimenter presented the appearance form of one object to the child and asked: “What do you think this is?” Then the real function was shown with the question: “But look, what is it really?” The procedure was then repeated with the child answering both questions again. Three fillers, such as “Have you ever seen such an object?,” were included to adapt the present training to Study 1 in length, without referring to another person’s perspective. The two training conditions differed in the following way: in the first-person condition,
children experienced their initial misconception of the objects during the training sessions and received appropriate feedback to their answers, whereas in the third-person condition, the children observed a doll being trained by the experimenter and assisted in presenting the nondeceptive objects to the doll. In order to minimize first-person experience in this training condition, the child was introduced to the appearance and the reality form of the deceptive objects at the beginning of the doll’s training unit. In this version, only the doll received feedback to its answers; the child’s own comments on the questions were ignored by the experimenter.

Results and Discussion
To evaluate the overall performance of children at a mean age of 3;0 in false-belief tasks, both training conditions were first combined to form one group of 48 children. Only children with correct control questions were given credit for a correct answer on a test question. Results on the Representational Change task showed a significant improvement in the self question, while children’s answers to the question about the other’s belief did not improve from pretest to posttest. In the Appearance-Reality task, 23% of the children gave correct answers to all three questions, and in the Change of Location task, 15% of the children gave correct answers. These values are roughly equivalent to those found for children in this age range without any training experience (Wellman, Cross, & Watson, 2001).

As for differences between training conditions, in the Representational Change task, children in the first-person condition outperformed children in the third-person condition on the self question; however, there were no differences between groups on the question about the other’s belief. There were no significant differences in overall performance between children in the two conditions on the Appearance-Reality or Change of Location task. Interestingly, on the Representational Change task, it was found that children who succeeded on the question about the other’s belief also succeeded on the self question (Kappa = 0.38, p = 0.001).
Thus, there are two main findings from this study. First, the study is the first to explicitly compare training conditions with self and others’ experience separated in a training setting. A significant difference between training conditions was found for the self question in the Representational Change task, with the first-person condition outperforming the third-person condition. This result might suggest that having the experience of being deceived, with discourse about that deception, focuses children’s attention on beliefs. However, the difference between the two conditions might also have been influenced by the type of engagement during the training (see also Wellman et al., 2001). The third-person group mostly observed a doll being trained. To keep the children interested and engaged in the training of a third person, children of this group assisted the experimenter and were asked questions irrelevant to the training. This more “passive” engagement did not lead to results comparable to those from the first-person training.

Second, an interesting issue is the question of age. Children at the age of 3;0 who were trained in perspective-shifting discourse and sentential complement syntax showed between 15% and 23% correct performance in transfer tasks that are very different in type and linguistic form from the training tasks. Wellman et al. (2001) found that similar numbers of children in this age group passed without any prior training experience. With regard to sentential complement syntax, the present results show that, at an age at which German-speaking children have already mastered some complement sentence constructions, this knowledge is not sufficient to permit the children to apply it in solving specific cognitive tasks such as false-belief situations.

Language and Social Understanding: A Developmental Progression

Young children depend on their skills of intention reading—knowing important things about what others see, do, intend, and attend to—to acquire and use linguistic conventions from soon after their first birthdays. In doing this, they come to understand something of the intentional structure of language and the way it can provide different perspectives and
descriptions of things: the very same animal is a dog, an
animal, a pet, or a pest; the very same action is running,
fleeing, chasing, or exercising. They thus come to form, in a
way that other animals do not, perspectival cognitive
representations (Tomasello, 1999).

But, sometime after their second birthday, many children
begin to command linguistic skills advanced enough to enable
them to engage in more sophisticated discourse
interactions with a real give-and-take of perspectives, that is,
those involving not just the different perspectives implicit in
the use of linguistic symbols but the explicit perspectives that
interlocutors linguistically express toward each other in
propositions—sometimes concerning each other’s previously
expressed propositions. As children engage in such discourse,
they are constantly simulating the perspective of the other
person and relating that perspective to their own (Harris,
1996; Tomasello, 1999). There are several forms of discourse
that seem especially important in children’s coming to
understand that others have beliefs. One is disagreements and
misunderstandings in which one person expresses the view
that X is the case, and the other disputes this and claims that
Y is the case. Also important may be (a) misinterpretations, in
which the adult interprets the child’s utterance in a way that
she did not intend, and (b) clarification requests, in which the
child says something that the adult does not understand and
so the adult asks for clarification. These situations lead the
child to try to discern why the adult does not comprehend the
utterance—perhaps she did not hear it, perhaps she is not
familiar with this specific linguistic formulation, and so forth.
In all, it seems that these kinds of disagreements,
misunderstandings, and repairs are an extremely rich source
of information about how one’s own understanding of a
linguistically expressed perspective on a situation may differ
from that of others.

But perhaps of the greatest importance is reflective discourse
in which the adult and child comment on the ideas contained
in the discourse turn of the other (or the self). For example, a
child may make some statement, and then either her
interlocutor or she herself will make some evaluative comment
about that statement. Relatedly, during the period from age 2 to age 4, for most languages, children also master sentential complement constructions (propositional attitude constructions) in which the speaker symbolically indicates both a proposition and some epistemic or modal attitude toward it, all in one construction. These could also then be considered, in a sense, reflective discourse, but in this case the state of affairs and the attitude toward it are all bound up in a single representational format in the form of a syntactic construction (de Villiers & de Villiers, 2000). Syntactic constructions are nothing other than grammaticalized (compressed and automated) strings of discourse (Bybee, 2003; Givon, 1979, 1995), and so looser discourse interactions and tighter syntactic constructions are all a part of the same process. It is thus natural that some kinds of syntactic constructions—specifically those that automate and compress reflective discourse—would be especially helpful in encouraging children to see both a state of affairs and a psychological attitude toward it all in one glance, as it were.

There is thus no conflict, in our view, between the view that discourse is crucial for developing an understanding of false beliefs and the view that mastering propositional attitude constructions is crucial. The recent work of Perner and colleagues (2003; see also this volume, chapter 11) demonstrates that mastery of the syntax of prepositional attitude constructions per se is not sufficient to engender false-belief understanding—since young German-speaking children master such constructions well before they pass false-belief tasks. However, the training study reported here shows that for children who are already on the cusp of understanding false beliefs, training with propositional attitude constructions is sufficient. We propose to think of these constructions as simply grammaticized discourse that, by being compressed and automated, contributes in special ways to a more general process of discourse in which children are constantly confronted with different perspectives and attitudes on different things, including their own perspective. Discourse is everything, but grammaticized discourse, that is, syntactic constructions, may influence children’s cognition in special ways.
Certain forms of social cognition—reading communicative intentions—are thus prerequisite for the acquisition and use of linguistic conventions, and the use of linguistic conventions, including syntactic constructions, in discourse is prerequisite for certain other forms of social cognition—understanding false beliefs. There is nothing odd about this; it is simply another manifestation of the dialectic in which children are biologically prepared for culture, but it is participation in culture—whose artifacts embody the cognitive skills and attitudes of past members of the culture—that takes their cognitive skills to new places (Tomasello, 1999; Vygotsky, 1978). To be precise about how all of this takes place, we must break down both language and theory of mind into more basic components.

References

Bibliography references:


presented at the 20tii Boston University Conference on Language Development, Somerville, MA.


