When we think about thinking:
The acquisition of belief verbs ☆

Anna Papafragou [a,*], Kimberly Cassidy [b], Lila Gleitman [c,d]

[a] *Department of Psychology, University of Delaware, 109 Wolf Hall, Newark, DE 19716, USA*
[b] *Department of Psychology, Bryn Mawr College, 101 North Merion Avenue, Bryn Mawr, PA 19010, USA*
[c] *Institute for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut St, Philadelphia, PA 19104, USA*
[d] *Center for Cognitive Science, Rutgers University, Busch Campus, Room A133, 152 Freylinghuysen Road, Piscataway, NJ 08854, USA*

Abstract

Mental-content verbs such as *think*, *believe*, *imagine* and *hope* seem to pose special problems for the young language learner. One possible explanation for these difficulties is that the concepts that these verbs express are hard to grasp and therefore their acquisition must await relevant conceptual development. According to a different, perhaps complementary, proposal, a major contributor to the difficulty of these items lies with the informational requirements for identifying them from the contexts in which they appear. The experiments reported here explore the implications of these proposals by investigating the contribution of observational and linguistic cues to the acquisition of mental predicate vocabulary. We first demonstrate that particular observed situations can be helpful in prompting reference to mental contents, specifically, contexts that include a salient and/or unusual mental state such as a false belief. We then compare the potency of such observational support to the reliability of alternate or concomitant syntactic information (e.g., sentential complementation) in tasks where both children and adults are asked to hypothesize the meaning of novel verbs. The findings support the efficacy of false belief situations for increasing the saliency of mental state descriptions, but also show that

---

☆ This manuscript was accepted under the editorship of Jacques Mehler.

* Corresponding author.

E-mail addresses: papafragou@psych.udel.edu (A. Papafragou), kcassidy@brynmawr.edu (K. Cassidy), gleitman@psych.upenn.edu (L. Gleitman).

syntactic information is a more reliable indicator of mentalistic interpretations than even the most cooperative contextual cues. Moreover, when syntactic and observational information sources converge, both children and simulated adult learners are vastly more likely to build conjectures involving mental verbs. This is consistent with a multiple-cue constraint satisfaction view of vocabulary acquisition. Taken together, our findings support the position that the informational demands of mapping, rather than age-related cognitive deficiency, can bear much of the explanatory burden for the learning problems posed by abstract words.

> – many of our complex ideas never had impressions that corresponded to them, and …many of our complex impressions never are exactly copied in ideas.
>
> David Hume (1739–40, 1.1.1. p. 3)

1. Introduction

Among the more awesome feats of learning in the human repertoire is the acquisition of the vocabulary of a first language. Analysis from Carey (1982) informs us that the average rate of word acquisition is something like 8–10 words a day though there are some hills and valleys along the way (Bloom, 2002). Furthermore, the consensus across individuals as to the interpretation of novel words is quite remarkable given the variability of input conditions, individual and population differences among learners and among languages learned (for reviews, see Gleitman & Newport, 1996; Waxman & Lidz, 2006). In the present article, we examine the acquisition of a particularly interesting class of words, the *belief predicates*, including such items as *think*, *believe*, *remember* and *know*. As we will discuss further, these items differ in their semantic-conceptual properties and in their typical learning signatures from most other words in several ways: they do not refer to perceptually transparent properties of the reference world; they are quite insalient as interpretations of the gist of scenes; they appear frequently in maternal speech to babies and yet occur in the child's own speech comparatively late; the concepts that they encode are evidently quite complex or abstract; and they are hard to identify from context even by adults who understand their meanings. It is immediately clear that the several particulars just mentioned are not independent of each other but are causally related in some ways that we will be trying to sharpen in the experimental work that is presented here. Our posit is that some fundamental aspects of the word-learning task can be uncovered by focusing attention on these unusual terms.

1.1. The learning trajectory for mental verbs

Belief verbs are notoriously hard to acquire. Children often produce verbs describing actions or physical motion such as *throw* or *run* before the second birthday

(Bloom, Lightblown, & Hood, 1975), and appear to understand them well (Huttenlocher, Smiley, & Charney, 1983; Gentner, 1978). *Think* and *know* do not usually appear until the end of the third year of life (Bretherton & Beeghly, 1982; Shatz, Wellman, & Silber, 1983) and are not well distinguished from one another in comprehension until around age four (Johnson & Maratsos, 1977; Moore, Bryant, & Furrow, 1989; Moore & Furrow, 1991; Naigles, 2000).

Part of the explanation for these differences in onset doubtless has to do with characteristics of the input speech to very young children. For example, owing to the greater freedom of nominal ellipsis in, say, Mandarin and Korean as opposed to English, verbs as a class account for a larger percentage of words uttered to children in those languages than they do in English (Tardiff, Shatz, & Naigles, 1997), and are more likely to occur in the perceptually prominent final position (Au, Dapretto, & Song, 1994). Moreover, Mandarin-speaking caretakers utter proportionally fewer abstract verbs to their infant offspring than do middle-class American caretakers (Snedeker & Li, 2000) and the transparency of contextual support accompanying use of such items also differs cross-linguistically (Snedeker, Li, & Yuan, 2003). All these factors influence both the overall emergence of verbs as a class in child speech, and the appearance of belief verbs.

However, these input differences are very far from providing the full explanation for the late appearance of belief verbs. Two major approaches, discussed in turn below, have been offered in the recent literature as providing fundamental explanations for these timing effects. The first, *the conceptual-growth hypothesis*, puts the spotlight squarely on the abstractness of the concepts that the belief terms express. The second, *the information-growth hypothesis*, does the same but asserts that the influence of abstractness is indirect: according to this second approach, it is not so much the conceptual difficulty of the concepts encoded that makes the trouble for lexical acquisition, but the difficulty for learners in discovering the conditions for these words' utterance during the flow of events. It is their insalience in the stream of events more than their (punct!) utter unavailability to thought that makes the trouble.[1]

1.1.1. The conceptual growth hypothesis

Plausibly, children's difficulties in acquiring the belief verbs are attributable to deficiencies with the underlying mental-state concepts (Bartsch & Wellman, 1995; Gopnik & Meltzoff, 1997; Perner, 1991). There may be a time during infancy (how early is a matter of controversy) when children can neither represent the ideas that these words express nor reason about them. Lacking the concepts, their verbal expressions cannot be acquired. Gopnik and Meltzoff (1997) present this view as follows:

> …the emergence of belief words like "know" and "think" during the fourth
> year of life, after "see," is well established. In this case…, changes in the
> children's spontaneous extensions of these terms parallel changes in their

---

[1] These issues resurface in the acquisition of other types of abstract/mentalistic vocabulary such as modal and evidential terms; for a recent discussion in the context of evidentiality (the linguistic encoding of information source), see Papafragou, Li, Choi, and Han (in press).

predictions and explanations. The developing theory of mind is apparent both in semantic change and in conceptual change. (p. 121).

This proposal about mental verbs is consistent with a broader view of word learning, according to which vocabulary development is an index of conceptual development (e.g., Dromi, 1987; Huttenlocher et al., 1983). The changing character of early vocabularies more or less directly reflects the growing typology of concepts in the mind of the learner. As Smiley and Huttenlocher (1995) put this:

> …Even a very few uses may enable the child to learn words if a particular concept is accessible. Conversely, even highly frequent and salient words may not be learned if the child is not yet capable of forming the concepts they encode…cases in which effects of input frequency and salience are weak suggest that conceptual development exerts strong enabling or limiting effects, respectively, on which words are acquired. (p. 20)

In the specific case of mental terms, the conceptual growth account seems to receive support from a very large body of experimental evidence showing that young children have difficulty with mental state reasoning (or *theory of mind*, cf. Premack & Woodruff, 1978). They appear to lack knowledge that others may be in a state of *false belief* and they fail to make inferences based on such an understanding (Wimmer & Perner, 1983; Carruthers & Smith, 1996; but see Leslie, 1994 for evidence against a conceptual deficit account and in support of the view that that the concept of belief is present in preschool children). Because the relevant thoughts are hard to come by early in life, words expressing them are accordingly hard to acquire.

1.1.2. The information growth hypothesis

An alternative, or additional, perspective on how mental verbs make their way into vocabularies has to do with the kind of evidence required for matching linguistic labels with mentalistic meanings (*the mapping problem*; cf. Locke, 1690). According to this view, the difficulty posed by the mental-content items is not solely or even primarily due to conceptual limitations in the learner; rather, the critical issue concerns the informational conditions inherent in the circumstances of early word learning and the learning tools that novices have at their disposal (Gleitman, 1990).

The idea here is that words that refer to mental states and events lack obvious and stable observational correlates: as a general rule, it is easier to observe that jumpers are jumping than that thinkers are thinking. Assuming that word learning – especially in its initial stages – relies heavily on establishing relations between what is said and what is (observably) happening in the extralinguistic environment, it follows that there will be barriers to acquiring mentalistic terms even where the concepts themselves fall within the learner's repertoire. In essence this is because, even though the relevance conditions for uttering *think* surely include thinking situations, this thinking takes place invisibly, inside nontransparent heads. These limitations would apply for any novice (even, say, a learned French-speaking professor visiting the USA on sabbatical) as long as he or she is acquiring vocabulary solely by observing how words line up with ambient objects, properties, relations, states, and events. This

learning procedure, which we have dubbed *word-to-world pairing*, has a venerable theoretical history and lies at the heart of associationist approaches to language acquisition (from Hume, 1739–40 to his many descendents in present day cognitive science, e.g., Barsalou, 1999; Smith, 2000).

As we will discuss further, the information growth hypothesis is that these barriers are breached by bringing new kinds of evidence to bear on the mapping problem. These additional sources of evidence are primarily linguistic, enabling a new kind of pattern search, *structure-to-world pairing* (sometimes called "syntactic bootstrapping").

1.2. A proposal for how mental verbs are acquired

In our view, both the approaches just sketched have something to say about why belief verbs are acquired relatively late. After all, these hypotheses are much alike in assigning the special difficulty of acquiring belief verbs (and, in fact, abstract words in general) to their unobservability and insalience during ordinary conversational exchange.[2] But while these approaches share this central insight, they differ in their implications for how these difficulties are to be redressed such that novices eventually acquire these harder lexical items. In the remainder of this introduction we will discuss linguistic and situational factors that plausibly relate to these issues, and experimental procedures by which they can be separately and jointly investigated.

1.2.1. Structural contributions to belief verb learning

A key to acquiring abstract words is the construction of language-internal evidentiary sources during the first few years of life (for a fuller description of such machinery, see Gleitman, 2003; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Trueswell & Gleitman, 2004). One such source is a growing database of linguistic co-occurrence information by means of which the learner can narrow the referential domain of new items (for instance, the co-occurrence of *eat* with edibles and *drink* with potables; Harris, 1951; Resnick, 1996; and see Altmann & Kamide, 1999 for recent experimental evidence). Even more central in the context of understanding belief verb acquisition is the verb's required and preferred structural environments. We know that a verb's syntactic privileges (namely, argument type, in conjunction with argument number and position) can predict the semantic profile of the verb (Baker, 2001; Croft, 1990; Fillmore, 1968; Goldberg, 1995; Gruber, 1965; Jackendoff, 1990; Levin, 1993; McCawley, 1968; among others). There is by now considerable

---

[2] The distinction between "observable" (e.g., *throw* or *give*) and "unobservable" (e.g., *hope* or *know*) is subtle at best and close to impossible to bring to ground in any very precise way. For example, *giving* comes in multiple flavors depending on the physical capabilities and manner of motion characteristics of the giving entity, so it is hard even to think about what is standardly observable when *giving* is uttered. Moreover, even this homely common term includes plenty of mental, hence unobservable, baggage; e.g., under most circumstances both giver and receiver must be sentient entities (while maybe you can get some blood from a stone, you cannot give it any). Volumes have been, continue to be and should be written on this topic. Here we mean only to allude to the gross generalization that in canonical contexts called "giving" there is potentially an exchange of goods out there in the world that, corrected for human perception and conception, can be seen or felt; whereas for "thinking," there decidedly is not.

evidence that such syntactic signatures provide a source of information that systematically cross-classifies the set of verbs within and across languages along lines of broad semantic similarity (Fisher, Gleitman, & Gleitman, 1991; and for cross-linguistic evidence, Geyer, 1994; Lederer, Gleitman, & Gleitman, 1995; Li, 1994). Furthermore, children are richly sensitive to these regularities in syntax-to-semantics mappings (Brown, 1957; Fisher, Hall, Rakowitz, & Gleitman, 1994; Gropen, Pinker, Hollander, Goldberg, & Wilson, 1989; Landau & Gleitman, 1985; Lidz, Gleitman, & Gleitman, 2003; Mintz & Gleitman, 2002; Naigles, 1990; Naigles, Gleitman, & Gleitman, 1992; Naigles & Kako, 1993; Pinker, 1989; Snedeker, Thorpe, & Trueswell, 2001; Waxman, 1994).

One might wonder how novices, especially small children, could break into these systematic correlations between form and meaning. Indeed if the correlations were arbitrary, as some have argued (e.g., Brown, 1957; Goldberg, 2004), they could play only a negligible and late role in learning. But on the contrary this interface is often remarkably straightforward, and the same form-to-meaning pairings manifest themselves again and again in the languages of the world, no matter how disparate they may be in other regards. For instance, thematic roles tend to line up one-to-one with argument positions in syntactic structures (roughly, the "theta criterion," Chomsky, 1981; see also Fisher, 1996; Fisher et al., 1994; Fisher & Gleitman, 2002; Lidz & Gleitman, 2004) and, in most languages, occupy distinctive structural positions. These features are reproduced, as well, in the self-invented languages of the isolated deaf (Gleitman & Newport, 1996; Goldin-Meadow, 2003; Senghas, Coppola, Newport, & Supalla, 1997).

Particularly pertinent in the present context, the lexical and phrasal composition of the arguments of mental verbs is quite transparently related to the logic of mental predicates. Sentential complementation ("He thinks *that it's red*") has semantic correlates: it implies a thematic relation between an animate entity and a proposition. This relation holds in every known language of the world and is a crucial feature of language design. Because of this restriction, the class of items licensing sentence complements includes verbs of communication (*say*, *tell*, *announce*; Zwicky, 1971), perception (*see*, *hear*, *perceive*; Gruber, 1965) and mental act or state (*believe*, *think*, *know*; Vendler, 1972). This syntactic behavior, to the extent that children can identify it, is a useful (and principled) cue to a novel verb's meaning. Strikingly, this structure–semantics correlation is implicated in blind as well as sighted children's discovery that *look* and *see* are terms of perceptual exploration whereas *touch* is not.[3]

---

[3] Notice that, for the blind, both touching (contacting manually) and seeing (apprehending via perceptual exploration) occur in the same situational contexts – a blind person must be in physical contact with an object etc., if he or she is to explore it perceptually. But the conditions for use of these terms are well-differentiated syntactically (one can "see that there's cheese in the refrigerator" but not "touch that the cheese is there"; Landau & Gleitman, 1985). Several authors have documented such correlations between the acquisition of mental verbs and competence with the linguistics of sentence complementation in young children, but taking varying positions as to the causal direction in which this correlation plays out in vocabulary acquisition (Hale & Tager-Flusberg, 2003; de Villiers & de Villiers, 2000).

Even though available for all kinds of lexical learning, syntactic information should be particularly helpful for the acquisition of mental verbs. This is for two main reasons. The first is the one we have emphasized most so far: to the extent that mental contents are not open to direct inspection (one may smile and smile and be a villain), the burden of learning falls the more heavily on internal linguistic-evidentiary sources. And second, mental verbs often take clausal complements that are semantically restrictive and hence quite informative about the kind of verb that can appear with them. The same two factors, but turned on their heads, apply to concrete action terms: They are more amenable to learning by situational observation, and their syntactic-constructional correlates are weak and undifferentiated within and across the class. They occur primarily in transitive or simple intransitive frames, which are themselves associated with a broad range of verb meanings; consequently, these frames provide little constraint on the kind of verb that can appear in them. For these kinds of verb, extra-linguistic information is the more useful in breaking into their meanings (Snedeker & Gleitman, 2004; Gleitman et al., 2005). Our first experimental prediction follows from these considerations:

> (i) The appearance of a novel verb in a sentence complement construction should narrow the search space for its meaning to a set including mental predicates – no matter what else is happening in the observable reference world.

1.2.2. Situational supports for belief verb learning

Acknowledging the usefulness of syntactic cues, the fact is that children do not learn the words *think* and *know* by inspecting incoming structures while being locked away from the world of objects and events. Despite the indeterminacy of ordinary cross-situational observation, we must believe that as a practical matter certain extra-linguistic circumstances and conversational settings facilitate thinking about thinking itself.

Consider in this regard one of the challenges facing the learner – the task of inferring that an observed discourse situation concerns *wanting* or *thinking*. This task is far from trivial: although children show an early sensitivity to cues about other people's mental states (including eye gaze, gestures, etc.; Bruner, 1974; Baldwin, 1993; Tomasello & Kruger, 1986), it is not clear how they become able to use that ability with the precision necessary to conjecture the meaning of a specific lexical item. After all, while there may well be jumping in view when one hears talk of jumping, there is usually no Rodin statue around when the child hears talk of thinking. Worse, there is generally some physical act occurring in the world or readily implied when one first hears a mental term uttered (for instance, the act of cupping one's chin in one's hand), and it is all too easy to infer, falsely, that this action is what the verb expresses. These "inferences to the concrete act" are often the rule in laboratory situations, where even adults are mulishly disinclined to suppose that the mental state of speaker or listener is under discussion. As one example, they reliably guess that a mother is uttering the verb *eat* when they see a (silenced) videotape of her handing a cookie to her young child. Actually, the mother is much more often alluding to the child's

mental state, saying "Do you *want* a cookie?" (Gillette, Gleitman, Gleitman, & Lederer, 1999). Evidently to learn that the sound /wônt/ encodes an abstract mental predicate, one must construct a quite global and abstract representation of the discourse gist. Under what circumstances is this likely to happen? When do we think about *thinking*?

For a mental state to emerge as a candidate for the meaning of a novel word, it must acquire some special conversational relevance. Otherwise, as in the example just given, the fact that there is a mental state (of desire) playing a causal role in the offering and consuming of cookies runs along under the attentional radar as a background supposition, while concentration falls on the actual (or even potential) physical act. In short, eating is "salient." That it is a desire-to-eat that gets you to eat is taken for granted. The hypothesis that we will be pursuing is the following:

> (ii) Mental states rise to special relevance in contexts where an incorrect mental state – a false belief – is easily inferable, and is at stake in the action observed.

This is likely to happen where a lie, trick or mistake is happening (Little Red Riding Hood *thinks* that this toothy hairy creature is her grandmother), or where there is uncertainty, contrast or change in the underlying mental contents ("What do you *think*?", "I used to *think...*"). The idea is that, much more than in the ordinary circumstances of true beliefs, circumstances involving false beliefs highlight the salience of the mental state itself. And only once the state of mind comes into focus can it support the relevant inference about the meaning of a novel word then uttered.[4]

Symmetrically, we predict no or little enhancement of the salience of mental predication during the vastly more common situations in which the speaker reports on a true belief or state of mind ("I think we should pick up the toys now"). These uses have been called *conversational* and have been broadly recognized to be uninformative for the language learner trying to figure out the meanings of mental verbs (Naigles, 2000; Shatz et al., 1983; Urmson, 1963). After all, what usually happens after or during such an utterance from caregiver to child is that the former (less likely, the latter too) begins picking up toys. It is the picking-up that relates transparently to the passing scene whereas the information value added by "I think …" in these true belief contexts is close to zero; if anything it is a politeness marker rather than a report of present mental contents, as Shatz et al. (1983) have pointed out. On the other hand, contrastive (or flatly false) uses of belief terms are potentially more informative. Studies tracing children's spontaneous uses of mental state verbs (Shatz et al., 1983) have treated contrastive uses of mental verbs as the first genuine indication of understanding their meaning ("Before I thought this was a crocodile; now I know it's

---

[4] This idea is consistent with recent proposals by Leslie, German, and Polizzi (2005) according to which, in verbal tasks, people tend to attribute true beliefs to others as a default (see also Leslie, Friedman, & German, 2004, for supporting data). The attribution of false beliefs is an exception to the rule, a non-default, and as such, is expected to be marked, or remarked upon, during conversation, as is typically the case with unexpected, or non-default, information (cf. Grice, 1967; Sperber & Wilson, 1986; see also Brown & Dell, 1987; Papafragou, Massey, & Gleitman, 2006, for discussion of related effects).

an alligator"). Our claim here is that such uses highlight the contextual conditions that promote discovery of the meaning of mental verbs. In short, from a learning perspective, we posit that *it is not the truth that sets you free*.[5]

1.3. A procedure for investigating word learning: the Human Simulation Paradigm

We have sketched two hypotheses about the learnability challenges posed by mental-content verbs. We take them as cooperative and complementary rather than competing. Both arise from the observational opacity of these words as they relate to the events, states, and scenes that accompany the linguistic input. So as to understand the machinery that supports lexical learning, it is nevertheless crucial to distinguish the effects of these situational and linguistic factors, and to examine their interactions. The procedure we will adapt in the experiments of Section 2 is a laboratory method known as the Human Simulation Paradigm (henceforth HSP; Gillette et al., 1999; Snedeker & Gleitman, 2004). This method investigates the potency of different kinds of cue (situational and linguistic) for solving the mapping problem for vocabulary growth. In the original HSP studies (Gillette et al., 1999), adults were shown silenced videos (each about 3–4 s in length) of naturally occurring scenes in which mothers were interacting and speaking with their infants aged between 18 and 24 months of age. These videos were selected to contain instances of the use of words that were highly frequent in maternal speech, as culled from a larger sample of naturalistic mother–infant interactions. At the exact point in the otherwise silenced video where the mother had actually uttered the target word, the participants heard a beep or a nonsense word. Subjects saw six such (randomly selected) videos for each target, and then guessed at its meaning. For example, one of these frequently used verbs was *call*. Not surprisingly, in the random selection of six scenes in which a mother was uttering this word to her infant, several of them showed either the mother or the infant herself playing with a toy telephone (and one or two did not, as even talk to babies often diverges from the here and now, or is opaque with respect to the context). Other participants were given linguistic information instead of or in addition to the videotaped situational information. This was in the form of the syntactic frame of the mother's actual utterance, with all the content words replaced by nonsense items. Consider again the verb *call*. One mother's utterance to her child had been "Why don't you call

---

[5] On the face of it, the hypothesis that false belief contexts are facilitative for child word learning may seem implausible, even paradoxical. After all, we know from a substantial body of experimental evidence that young children have enormous difficulty understanding that others may be in a state of false belief and of making inferences based on this understanding (Wimmer & Perner, 1983; cf. Section 1.2.1). Doubtless such difficulties are related in some ways to late appearance of the belief terms. However, from another perspective, it might not be so surprising after all if false belief situations did turn out to capture the attention of very young children under certain circumstances. There has been considerable controversy as to the extent to whether failures by 3- and 4-year-olds in false belief tasks actually represent conceptual inadequacy as opposed to difficulties with the kinds of explicit judgmental task ordinarily used for testing (see Leslie, 2005; Bloom & German, 2000, and many others). Indeed, recent experimental findings suggest that sensitivity to the true belief/false belief distinction can be demonstrated in infants under 2 years if the task is made implicit enough (Onishi & Baillargeon, 2005).

Daddy?". For this item in the syntax-information condition, the adult participant was shown the nonsense-replaced version "Why don't *ver GORP telfa*?". Another of the maternal utterances was "I'm gonna call Markie" and for this the participant was shown "*Mek* gonna *GORP litch*"; etc. Finally, a third set of participants watched the videotapes while also being shown their matching nonsense-containing syntactic frames, that is, a scene plus syntax combination.

Two results from such prior experimentation are particularly relevant for our present purposes. The first has to do with the varying usefulness of observational versus syntactic cues, depending on the verb type. Situational evidence taken alone led to some success when the targets were concrete action terms such as *push* and *catch* (35–50% accuracy) but was an especially poor cue for mental expressions such as *want*, *know*, *like* and *think* (0% accuracy). The findings in the syntactic condition of this experiment were strikingly reversed. Syntactic frames provided particularly strong support for the acquisition of mental verbs (90% accuracy) but were far less helpful in the case of concrete verbs (40% accuracy). In short, if observation is not informative then the information has to come from elsewhere; here, from syntactic context.

It remains to be noted, however, that the videotaped scenes of mother–child interaction that subjects in these experiments observed were not well designed to elicit conjectures about belief verb use; that is, they were not discourse settings of trickery, oddity, or false belief. In the present experiments we will try to put the world back into the learning equation for these hard words by introducing them in the kind of situational context that we conjecture to be especially useful in amplifying their salience: the false belief situation. In a related regard, one further result of interest emerged from the original HSP findings. This has to do with the power of converging cues. Experimental participants who received both situational (the videotaped scenes) and syntactic (the nonsense-containing frames) information about the mystery words were close to ceiling in identifying both action and mental-content verbs (approximately 80% success rate overall; see particularly Snedeker & Gleitman, 2004 for a replication that systematically varied and combined several cue types).

1.4. Summary, and an experimental prospectus

Our group of investigators has claimed that the kinds of finding just summarized *for adults* are suggestive as well for why the very young child's vocabulary has the character that it does. The child is predicted to be delayed in the acquisition of mental-content terms owing primarily to the lack of linguistic resources that are the best cues to the identification of such items. As the child builds up distributional and syntactic knowledge of the exposure language, these linguistic cues contribute to the learning of abstract vocabulary items, the belief verbs in particular. Of course, it is never denied that conceptual changes may exert collateral influences in the same directions, though not measured in these prior experiments.

In the studies now to be presented, we set out to identify the contribution of syntactic cues (e.g., sentential complementation) and potentially helpful cues from observation (the presence of a false belief) to the identification of belief predicates.

2. Structural and situational cues to the belief vocabulary: experimental approaches

In this section, we present three experiments. The first two, adapting HSP methods, are designed to understand conditions under which adults are likely to utter terms that express mental acts and states, particularly beliefs. This probe into the conditions of adult utterance has *prima facie* relevance to the learning conditions for young children in this same regard: you cannot acquire the word if your mother does not utter it. A subsidiary aim of these adult-participant experiments is to establish a baseline for the mapping of these words when the simulated learners are cognitively mature; this bears on the cognitive growth versus information growth accounts of vocabulary learning, as described in Section 1. Experiment 3 turns at last to preschool children who are in the first major verb-learning period of life, also testing adults for control and comparison purposes. This last experiment is designed to examine the extent to which the conditions under which adults speak are also the ideal ones under which children can acquire the belief terms.

2.1. Experiment 1: eliciting mental verb utterance in false belief contexts

This experiment was designed to inquire how a linguistically competent and cognitively mature population interprets situations of action, desire, and belief as cues to describe these in mentalistic terms. At their strongest, such results could point to the input conditions for belief verb use by caregivers. Of course, creating conditions that increase salience and thus usage probability for the adult caregivers may not, *ipso facto*, have a like effect on their watching–listening offspring. Nevertheless, a first step in finding out how the mentalistic verbs are acquired is discovering the conditions under which they are uttered by the adult community.

2.1.1. Participants

Participants were 96 undergraduate students from the University of Pennsylvania and Bryn Mawr College. Students were enrolled in introductory psychology classes and received course credit for participation.

2.1.2. Materials

Brief silent scenarios showing actors engaged in simple activities were prepared and videotaped. These scenarios were designed to fall into three broad event/state classes: belief scenes, action scenes, and desire scenes. In detail:

Belief scenarios (B) depicted either an agent who was tricked or mistaken (false belief scenes, henceforth FB), or almost tricked or mistaken (true belief scenes, TB).[6] For each of four belief scenarios, we actually constructed three variants: one TB version and two FB versions which varied in how far removed from reality or the particulars of the depicted event the false belief was. Thus there were 4 sets of belief scenes, all 3 variants within a given set showing the same event but with a different

---

[6] The false belief scenes were inspired by videotaped events developed by Jill de Villiers for the study of false belief understanding in children. We thank her for making her materials available to us.

denouement (thus, 12 belief videos altogether). The following is an example of one such set:

> A woman is sitting at her kitchen table drinking tea and reading the paper. On the table there is a flower vase, a teapot and the woman's teacup. Another woman walks in and begins wiping up the table and moves things around a bit. The first woman does not see or notice the changed position of the three pertinent objects. Rather, she simply reaches for her coffee cup while still reading. The objects are now positioned close enough to each other that such an absent-minded reach may result in the right or wrong item being selected. Each variant videotape now depicts a different outcome:

> *True Belief* (*TB*) *scene*: in this variant the woman is lucky. Her hand fortuitously lands on the teacup, which she picks up and brings to her mouth.

> *Plain Vanilla False Belief* (*PFB*) *scene*: the woman accidentally picks up the teapot and does not notice her mistake until the teapot is close to her mouth.

> *Zany False Belief* (*ZFB*) *scene*: the woman accidentally picks up the flower vase and does not notice her mistake until the vase is close to her mouth.

> The difference between PFB and ZFB scenes has to do with whether the false choice is one that has some appropriateness to the woman's implicit goal (after all, there is tea in the teapot) or is risible (there is only water and flowers in your prototypical flower vase).

Action scenarios (A) showed an agent performing some common action (e.g., a girl drawing a picture). There were 4 such scenarios.

Desire scenarios (D) showed an agent trying to do something or looking for something (e.g., a girl was trying to reach a box on a high shelf). There were 4 such scenarios.

2.1.3. Procedure

Participants watched a series of 12 of the video clips (all 4 A, all 4 D, and one variant of the 4 B scenarios) in individual sessions. To control for effects of order and item choice, three basic lists were created, each containing one variant for each belief scene, which were then reversed to create 6 different stimulus orders; the result was 6 stimulus orders. Participants were randomly assigned to these order lists. They were told that the clips had been designed to be shown to mothers and their young children and that they should guess what the mothers might say to their children when watching these videos together. Participants were asked also to provide a second response in each case ("What else do you think the mother might have said in this situation?"). After each clip the video was stopped to allow participants time to write down their answers. Each session lasted about 20 min.

2.1.4. Predictions

Based on previous studies using similar paradigms (Gillette et al., 1999), we expected participants to conjecture action verbs (A) more often than other categories

across all scene types. Crucially, however, we expected the distribution of verb types to differ systematically as a function of scene types. A critical prediction was that the proportion of belief (B) verbs would increase in FB scenes to the extent that these violations of expectation really increase the observer's attention to actor's state-of-mind. For related reasons, we predicted that this effect would be amplified by the striking oddities in the ZFB variants. Finally, we predicted that if participants were given a second chance to describe the scene, their responses would show more B responses than their first guess, as the latter could have satisfied their need to describe the scene in terms of its concrete action content.

2.1.5. Coding

Verbs in participants' responses were classified individually as A (Action: *read*, *move*), B (Belief: e.g., *decide* or *notice*), D (Desire: e.g., *wish* or *try*), or O (Other: e.g., *let* or *have*). (See Appendix 1 for the full list). Participants' total responses per guess typically included more than one clause and hence more than one verb. Sample responses in the (first) guess for the PFB teacup scene were (verbs italicized):

> The girl almost *drank* from the pot because she did not *realize* her mug was *moved*. She does not *know* her cup has been *moved*, so she *thinks* it's in the same place as before.

The coding in the first instance would include a (single) entry in 2 categories because the response contained at least one member of the A category (*drink*, *move*), and of the B category (*realize*). Similarly, the second sentence would be coded as A (*move*), B (*know*, *think*) and O (*be*).[7]

2.1.6. Results

The results were submitted to preliminary analyses to determine if presentation order and guess type had influenced the response styles. Each participant was assigned two scores based on the proportion of B responses he or she offered as a first and second guess for the belief (TB, PFB, and ZFB) scenes. Those scores were entered into an ANOVA with Order (1, 2, 3, 4, 5, 6) as a between-subjects factor and Guess (1st, 2nd) as a within-subjects factor. The analysis returned no significant main effect of Order and no interaction between Order and Guess; for all subsequent analyses, we ignore Order as a factor. The analysis also revealed no significant main effect of Guess: participants produced B verbs on average 27% of the time in both their first and second guesses. Since, despite our predictions, this "second chance guess" yielded results so eerily identical to the first chance guesses, further analyses collapse across the two guesses (thus doubling the number of codable responses).

---

[7] As these examples show, each B verb in subjects' responses was typically accompanied by a sentential complement containing an A verb ("She doesn't *know* her cup has been *moved*"). Thus if we only coded verbs individually, we would inflate the proportion of A over B responses. Our coding scheme classifies responses on the basis of whether they contained at least one verb of a certain category, regardless of other verbs in the sentence.

2.1.6.1. Verb type and mental state. Table 1 shows the pertinent results (the proportion of Other verbs was just 4% of all verbs offered so we will not discuss them further). As is obvious by inspecting the Table, participants overwhelmingly often (over 90% of the time) included an A verb in their response, no matter the scene type. D and B scenes were also clearly effective in increasing the proportion of responses of their respective categories in the descriptions of the corresponding scenes: the scenarios that we intended as demonstrating attempts to accomplish something increased D responses to 37.79% of all answers, while in other scene types the number of D responses hovered around 10%. The scenes that were meant to depict mental state as a relevant part of interpretation increased B responses (from around or below 10% in the A and D scenes to approximately 40% of responses in the FB scenes combined). In particular, B responses increase threefold (17–41%) as between the TB and FB categories. All the same the absolute proportion of B verbs is lower than might have been expected, a finding we retraverse in discussion below.

We now turn to the statistical analysis of the results. As is clear from the Table, contrary to expectation, subjects' responses did not differ in the kinds of verbs they contained between "zany" and plain vanilla FB scenes (this was confirmed by pairwise comparisons). In all further analyses we thus combine the results for ZFB and PFB scenes.

An ANOVA using the proportion of B responses elicited for each item as the dependent variable with Scene (A, D, TB, FB) as a between-items factor revealed a significant main effect of Scene ($F(3, 16) = 48.64$, $p < .0001$). The efficacy of FB scenes over A and D scenes in eliciting B responses was confirmed through pairwise comparisons (FB vs. A: 41% vs. 6%, $t(10) = 9.203$, $p < .0001$; FB vs. D: 41% vs. 9%, $t(10) = 8.603$, $p < .0001$, respectively). TB scenes seem to have an intermediate status, as they elicited reliably more B responses (18%) than did either A scenes ($t(6) = 3.552$, $p = .01$) or D scenes ($t(6) = 2.85$, $p = .02$), but reliably fewer such responses compared to FB scenes ($t(10) = 5.66$, $p = .0002$).

A similar ANOVA using the proportion of A responses as the dependent variable and Scene as a between-items factor returned no main effect of Scene. The same type of analysis over the proportion of D responses also did not yield a significant main

Table 1
Proportion of responses (averaged over the two guesses) that include the three main different types of verbs (Exp. 1)

Scene	Verb category		
	Belief (B)	Desire (D)	Action (A)
Action (A)	6.44	9.86	92.88
Desire (D)	9.31	37.79	90.42
True Belief (TB)	16.96	10.43	93.04
False Belief (FB) total	41.00	9.54	90.46
Plain False Belief (PFB)	39.06	8.58	91.42
Zany False Belief (ZFB)	42.98	10.53	89.47

Note that because responses usually contained more than one verb item, the proportions add up to more than 100%

effect of Scene. Nevertheless, pairwise comparisons revealed a statistically significant difference in proportion of D responses when D scenes were compared to A scenes (37% vs. 9%, $t(6) = -6.67$, $p = .0005$) and TB scenes (37% vs. 13%, $t(6) = -3.72$, $p = .0098$), but not FB scenes (37% vs. 16%, $t(10) = .5$, $p = .62$).

2.1.7. Discussion

This study was our preliminary attempt to examine the role of visual-situational input in eliciting the utterance of mental-content verbs in a scene description task. As predicted, participants' responses overwhelmingly often (over 90% of the time) included action verbs, those that describe the unfolding event. Responses containing mental (belief or desire) predicates are comparatively less frequent, occurring in less than 20 percent of participant responses overall. Even in scenes specifically designed so that desire is highly relevant, desire states are reported only about 40% of the time. Similarly, in scenes designed so that belief contents are highly relevant, belief states are reported only about 20% of the time. Thus *acts trump motives* according to our adult respondents, at least under these constrained laboratory conditions. All the same, the false belief situations did raise the proportion of mental state reporting by a significant proportion, not only over their proportion in all scenarios (from 18% to 41%) but in their proportion compared to true belief situations (from 17% to 41%). If such conditions (in the real world) turn out to be those that elicit thoughts of state of mind in the adult speaker, perhaps they do so in the word-learning novice as well. To the extent so, false belief situations would in this way have created an environment specially tuned to the acquisition of mental predicates.

2.2. Experiment 2: a more stringent test – limiting the cues

Experiment 1 used silent videos, withholding all internal linguistic evidence as to the semantics of the nonsense items. In the present study, adult participants were asked to conjecture verb meanings under varying informational conditions, in an adaptation of the HSP. The question is how situational (action vs. belief) and syntactic (transitivity vs. S-complementation) evidentiary sources are used, singly and jointly, to solve the mapping task for belief verbs by a cognitively competent adult population.

2.2.1. Participants

Participants were 56 undergraduate students from Bryn Mawr College enrolled in introductory psychology classes; they received course credit for participation.

2.2.2. Materials and procedure

Given that there were no effects of stimulus list in Experiment 1, one list was randomly selected for use in this experiment. The list included 3 false belief (FB) scenes, 1 true belief (TB) scene, 4 action (A) scenes and 4 (D) desire scenes, as described previously.

Participants were randomly assigned to one of three information conditions. For all conditions, participants were told that they were going to try to guess what

particular verb mothers would say to their young children, as cued by the information given:

(a) *Scene Only Condition*: participants watched the video with no sound (as in Experiment 1). For example, one of the FB scenes was the one described in the *Materials* section of Experiment 1 as "the teacup scene."

(b) *Syntax Only Condition*: participants were presented with a (written) list of 12 sentences that referred to the actions and states of mind depicted in each of the scenes (though these participants never saw the scenes themselves). Following the HSP procedure (e.g., Snedeker & Gleitman, 2004), the sentences had been doctored, Jabberwocky style, by replacing each of its content lexical items with a nonsense word. The verb itself, also a nonce word, was capitalized to show that this was the item that participants had to identify. The list contained 4 transitive frames with a direct object (e.g., *Vamissa VAMS the torp*), 4 frames with clausal complements introduced by *that* (e.g., *Vamissa LODS that she ziptorks the siltap*) and 4 intransitive filler frames (e.g., *Vamissa TROMS*).

(c) *Scene+Syntax Condition*: this group of participants watched the same videos as in the Scene Only Condition, and received the written list of syntactic frames used in the Syntax Only Condition. The syntax and the scenes were paired such that the syntactic frames "matched" the scenes. Thus, the 4 B scenes were paired with clausal complement frames and the 4 A scenes were paired with transitive frames (the remaining filler D scenes were paired with intransitive frames). Participants looked at the syntactic frame, watched the scene and were asked to guess the verb using both pieces of information.

2.2.3. Predictions

In light of the results of Experiment 1, we expected that in the Scene Only Condition, action verbs would be the predominant response, but that A-bias would be greatly reduced for FB scenes. Similarly, we expected that sentence complement (S-comp) structures in the Syntax only condition would enhance mental verb responses. Finally, we expected an overadditive proportion of mental verb responses in the Scene+Syntax Condition when the stimulus was a FB scene paired with an S-comp structure.

2.2.4. Coding

The proportion of B verbs guessed for each critical Scene type (A, FB) and Syntactic type (Transitive, S-comp) was computed.

2.2.5. Results

The results for test items are summarized in Table 2. Inspection of the Table shows that for A scenes, B guesses were very low across information conditions (from 4% to 7%). For these A scenes it is important to keep in mind that the syntax given was the transitive construction. In contrast, for FB scenes, the proportion of B verb responses increased (from 5% to 60% overall). This increase in B-verb proportion was even larger (up to 78%) in the Scene+Syntax condition when S-comp syntax was paired with the FB scene. Thus, for the FB Items, information condition seemed to have an effect.

Table 2
Mean percentage of belief verb guesses per item type (Exp. 2)

Action (A) items	
Scene Only Condition	4.17 (9.60)
Syntax (transitive) Only Condition	6.58 (14.05)
Scene + Syntax Condition	3.95 (9.37)
Total	4.91 (11.10)
False Belief (FB) items	
Scene Only Condition	44.44 (32.46)
Syntax (S-comp) Only Condition	59.68 (36.22)
Scene + Syntax Condition	77.26 (34.95)
Total	60.75 (34.95)

The statistical analyses corroborated this pattern. Using the number of B verb conjectures as the dependent variable, we conducted a mixed model ANOVA with Scene Type (A, FB) as a between items variable and Information Condition (Scene Only, Syntax Only, Scene + Syntax) as a within item variable. The analysis resulted in a significant main effect of Scene Type ($F(1, 5) = 1,082.41$, $p < 0.001$), and a significant main effect of Condition ($F(2, 5) = 10.63$, $p < 0.005$). However, these main effects were qualified by a significant interaction between Item Type and Condition ($F(2, 5) = 14.65$, $p < .005$). As described above, for the A scenes, there were few cognitive state verb guesses and the number of guesses did not change as a function of condition (pairwise comparisons (Bonferroni), all p's > .6). In contrast, for the FB items, more B guesses were made in the Scene + Syntax Condition than in the Scene Only Condition (pairwise comparison (Bonferroni), $p < .05$). In the Syntax Only Condition B guesses were at an intermediate level for the FB items (S-comp frames). They did not differ significantly from either of the other conditions (pairwise comparisons (Bonferroni), all p's > .4). Thus, while overall more B guesses occurred for the FB items, information condition affected the number of B guesses for these items.

Table 3 shows the results presented by participants. From the table, it is evident that subjects in every condition made many more B guesses for the FB items than for the A items; however, this difference was greatest in the Scene + Syntax Condition because of the higher level of B guesses for the FB items in this Condition. Consistent with this pattern, a mixed-model ANOVA with Item type as the within subject variable and Condition as the between-subject variable revealed a main effect of Item type ($F(1, 53) = 176.01$, $p < .001$). There was also a main effect of Condition ($F(2, 53) = 3.43$, $p < .05$). These main effects were qualified by a significant Item type by Condition interaction ($F(2, 53) = 5.25$, $p < .01$). As can be seen in Table 3, for the A scenes, participants across all conditions were equally unlikely to guess cognitive state terms (pairwise comparisons (Bonferroni), all p's > .5). In contrast, for the FB items, participants made more B guesses in the Scene + Syntax Condition than in the Scene Only Condition (pairwise comparison (Bonferroni), $p < .05$). Participants in the Syntax Only Condition were in between

Table 3
Mean percentage of belief verb guesses by participants per condition (Exp. 2)

Scene Only Condition	
Action Scenes	6.50 (3.70)
False Belief Scenes	40.33 (3.79)
Total	21 (18.40)
Syntax Only Condition	
Transitive Frames	8 (1.40)
S-comp Frames	55 (12.53)
Total	28.14 (26.16)
Scene + Syntax Condition	
Action Scenes + Transitive Frames	3.75 (4.79)
False Belief Scenes + S-comp Frames	79.67 (8.62)
Total	36.29 (41.02)

participants in the Scene Only Condition and the Scene + Syntax Condition in their proportion of cognitive term guesses. They did not differ significantly from either of the other conditions for the FB items (pairwise comparisons (Bonferroni), all p's > .3).

2.2.6. Discussion

Once again when faced with the task of matching a scene with a potential verbal response (Scene Only Condition), participants overwhelmingly often (79% percent of the time) came up with action verb responses: their bias was to describe the unfolding event with its motives presupposed and unmentioned. However, in those cases where a character's behavior was odd owing (presumably) to a false belief, mental verb conjectures increased significantly (to 45%) In Scene + Syntax trials where observational (FB) and syntactic (S-comp) cues converged on a false belief state, the formerly insalient B interpretations increased to almost 80% of responses. As always, it is a long distance from simulations to the real world, but we believe the results link plausibly to the conditions that inspire cognitively mature individuals to think about scenarios in terms of their motives, and thus to utter belief verbs. But do these inspirations exist in much the same way for the child listeners as well? If so, then they too will be most likely to interpret novel verbs as belief verbs under the same linguistic and extralinguistic conditions. The next experiment attempts to investigate this question.

2.3. Experiment 3: what makes you think you are hearing a mental verb? Situational and linguistic contexts

Now we ask when children (and adult controls) think that a novel verb that they are hearing is describing a mental event or state. Overall, the experiment is designed to see how the syntactic and situational cues uncovered in Experiments 1 and 2 conspire to predict preschoolers' verb conjectures.

2.3.1. Participants

Participants included 34 children enrolled in Philadelphia daycares/preschools, ranging in age from 3;7 to 5;9 years (mean 4;5). Data from three more children were excluded from all analyses due to a failure to meet a pre-test threshold. An equal number of adults also participated. Most of these latter were undergraduate students enrolled in introductory psychology classes at the University of Pennsylvania who received course credit for their participation.

2.3.2. Procedure

Children were tested individually by a single experimenter. At the start of the experiment, children were invited to "watch some stories on the computer" with the experimenter. They were told that some of the people in the stories used funny words, words that the experimenter did not know. The experimenter then explained that children's help was needed to understand what the characters were trying to say. We know from prior work that children are very willing to offer their assistance in this sort of setting, as they consider themselves the true intermediaries between fictional characters and the adult world (e.g., Fisher et al., 1994; Waxman & Gelman, 1986).

Children went on to watch a series of videotaped stories. All stories followed a fairytale format with a pre-recorded narrative that described the observed events. This format was chosen to ensure that children would maintain their interest. At the end of each clip, the participants received a prompt from one of the scene's characters describing what happened in the scene; the prompt was always a sentence in which the verb was replaced by a nonsense word. The experimenter repeated the sentence and asked the child what the nonsense word meant. (Adult participants were also individually tested in the same way but were given a mercifully shorter introduction to the word-guessing game.)

Stories were administered in 4 pseudo-random orders. Each subject saw a total of 12 stories (Scene type/Syntax combinations). These included 4 desire scenes plus 8 belief scenes (2 of each of the 4 possible Belief type Frame combinations without repetition within a single belief set; see below for details).

Participants completed this task in two sessions interrupted by a short break. This scheduling was to ensure that even the youngest children would be able to complete the task. Overall, participants were successful in coming up with a verbal response: missing responses accounted only for 2.2% of the child data and 0.7% of the adult data. No child who had passed the pretest failed to complete the full sequence.

2.3.3. Materials

The critical test items consisted of a set of 8 belief scenes ranging in length from 33.07 s to 69.21 s ($M = 53.25$). Each belief scene had four variants which were created by fully crossing two factors: type of belief (true vs. false) and type of syntactic frame within which the novel verb was uttered (transitive frame with a direct object vs. clausal complement introduced by *that*). The following is an example of what participants heard and saw during a belief scene (and its variants):

Once upon a time there was a little boy named Matt who had a sick grandmother. One day his father says to him, "Take this basket of food to your sick grandmother; it will make her feel better."

False Belief (*FB*) *scene*:

So the little boy, Matt, says, "Okay." And so Matt walks to his grandmother's house through the spooky forest where he meets a big, bad cat."Where are you going with that basket of food?" asks the cat."To my sick grandmother," says Matt. The cat says to himself, "I will trick the boy so I can steal his food!" The cat runs to the grandmother's house before the little boy gets there. He locks her grandmother into the closet and hides under his grandmother's blankets. Matt knocks on the door and goes over to his grandmother's bed. "Hi, grandmother. Here is a basket of food to make you feel better."

[S-comp Frame] *Cat*: "Did you see that? Matt GORPS that his grandmother is under the covers!"

[Transitive Frame] *Cat*: "Did you see that? Matt GORPS a basket of food!"

True Belief (*TB*) *scene*:

So the little boy, Matt, says, "Okay." And so Matt walks to his grandmother's house through the spooky forest where he meets a big cat."Where are you going with that basket of food?" asks the cat."To my sick grandmother," says Matt. The cat says, "I am sorry your grandmother is sick. Please let me help you carry the basket of food to your grandmother." [Matt hands the basket to the cat and they continue to his grandmother's house]. Matt knocks on the door and goes over to his grandmother's bed. "Hi, grandmother. Here is a basket of food to make you feel better." [S-comp Frame] *Cat*: "Did you see that? Matt GORPS that his grandmother is under the covers!"

[Transitive Frame] *Cat*: "Did you see that? Matt GORPS a basket of food!"

The participants' task was to identify the meaning of GORP on the basis of the combination of observational and syntactic cues present within a single variant.

A total of 4 desire (D) scenes were used as fillers. These scenes were shorter than test stimuli, ranging in length from 9.14 s to 31.07 s ($M = 19.61$). In D scenes, a person was either trying to do something (e.g., reach a teddy bear), or expressing strong emotion over not being able to do something (e.g., getting out of a locked room). In such scenes, the target verb always appeared with an infinitival complement:

Desire (*D*) *scene*: Once upon a time there was a boy who liked to hug his teddy bear. But his teddy bear was way up high so the boy could not reach it. [The boy tries to reach the toy]

[Infinitival Frame] *Boy*: "I PLEEB to get my teddy bear."

In addition to the test items, 3 pretest video clips were constructed. These pretests were short (6 sec) clips of simple actions without narration, designed to identify children who did not grasp the general nature of the word-guessing task. For instance, in one of these clips, a boy was shown throwing a large ball in the air a number of times. At the end of the clip, the boy said: "I'm DOLTING". At this point, the experimenter repeated the sentence and asked the child what it meant. If the child responded nonsensically or provided no answer, the experimenter then showed the clip again and prompted for a response. Only children who passed at least two of the three pretests were allowed to complete the main phase of the experiment.

2.3.4. Predictions
On the basis of Experiments 1 and 2, we expected that the FB scenarios would render the actors' inferred states of mind more salient and thus would be significantly more likely than the TB scenarios to yield mentalistic interpretations. Additionally, based on previous studies (e.g., Gillette et al., 1999), we hypothesized that S-comp frames would promote belief verb conjectures. Our main hypothesis was that the two cues together (FB scenarios/S-comp verbal descriptions) would be especially supportive of mentalistic conjectures; that is, we expected that the identification procedure for vocabulary is a probabilistic multi-cue constraint satisfaction system (Gleitman et al., 2005). An important sub-hypothesis was that syntactic cues would largely overwhelm observational biases when these cues were in conflict (the FB scene/T frame and TB scene/S-comp frame conditions). This hypothesis was based on the almost categorical tendency for nonsense sentence-complement sentences to elicit cognitive verb conjectures in prior work (Snedeker & Gleitman, 2004) and in general the statistical though not categorical priority of linguistic cues for both children and adults (e.g., Fisher et al., 1994). A related prediction was that the patterning of child and adult responses would be approximately the same, consistent with the information change hypothesis rather than the developmental change hypothesis (as described in the Introduction).

2.3.5. Coding
Responses were coded for two properties:

(1) V*erb type*: individual verbs in participants' responses were coded under the categories action (A), belief (B), desire (D) and Other (O). For a complete list of verbs offered by the child and adult participants, see Appendix 2.
(2) S*trict frame compliance*: the verbs in participants' responses were also coded for whether they were legitimate ("grammatical" or "acceptable") in the syntactic frame of the query that was posed. So, for example, the response "He ate all" to the frame "Matt DAXES that the cookies are all gone" was coded as non-frame compliant because the verb *eat* does not accept complement clauses. This analysis is meant to capture even more fine-grained differences among verb syntactic privileges than our coding into 4 types reveals. Not all belief verbs occur in the kind of S-comp structure that we used as stimuli, namely cases in which the embedded sentence is introduced with *that* and

contains a finite verb ("I think that he is dancing"). For instance, verbs of questioning such as *wonder* do not (or "hardly do") occur with *that* complementizers; rather, with *if* or *whether* ("I wonder if/whether grandma found the basket"). Strictly frame compliant responses, then, were coded when the participant offered a verb which fit this more restrictive criterion. Notice also that a one word response such as "think" but also a multiple word response such as "He thinks the cookies are all gone" both can meet the criterion of frame compliance and were so coded.

2.3.6. Results

The effect of stimulus order was nonsignificant for both child and adult participants, so all further analyses collapse across order groups.

2.3.6.1. Actions speak louder as words. Figs. 1a and b show the verb response types for adults and children, respectively. What jumps out at the eye immediately is the high incidence of A verbs in both populations (31.60% and 43.90% of responses for adults and children, respectively). Even if we add together both types of mentalistic response (D and B types), they are still outnumbered by A responses for the children (35% vs. 44%, respectively) though not for the adults. Moreover, the belief verbs are especially dispreferred by both age groups (23.5% and 11.3% of all adult and child responses, respectively). There were 4 D scenes and 8 B scenes, so in a fair world we ought to see twice as many B verb responses as D verb responses, but that is not what happened: D responses are more than double the B responses for children (23.5% vs. 11.3%) and a difference in the same direction is nearly as dramatic for adults (30.9% vs. 23.5%).[8] These findings underline the point that the actions rather than the states of mind of the character are assumed under standard circumstances to be the topic of the conversation – that which is probably being spoken about, whether or not it is also what's being thought about.[9]

---

[8] We will not discuss desire verb (D) responses in detail (but see Tables 4a and 4b). For child participants, D verb conjectures are more frequent in scenes where the character is trying to do or get something than in (both types of) B scenes (33.1% vs. 6%) and the same is true for adult participants (86.8% vs. 3%). Nevertheless, even in D scenes, A verbs still account for the majority (59%) of children's responses; for adults A verbs only account for 11.8% of these responses.

[9] One might wonder whether general frequency facts about verb use affect participants' responses: certainly the belief verbs as a group are less frequent than some action verbs and it might be argued that for this reason alone they might come less readily to mind. Actually, whether frequency has any effect at all in this setting is in doubt. After all, the relevance and syntactic context of the verb ought to be the main determinant, with word frequency having its effect only within that limitation (e.g., *thinking* is expected rather than *hypothesizing*). Such distinctions do not arise given our coding scheme which divides the verbs only according to gross semantic class. However, for completeness, Appendix 2 includes a frequency analysis for the verbs offered by our participants both in terms of the overall set of data and in terms of the frequency of the words per million English words (Kučera & Francis, 1967). As the Appendix shows, desire verbs and certain belief verbs such as *think* or *know* are highly frequent in the corpus but are only infrequently offered as conjectures by our participants.
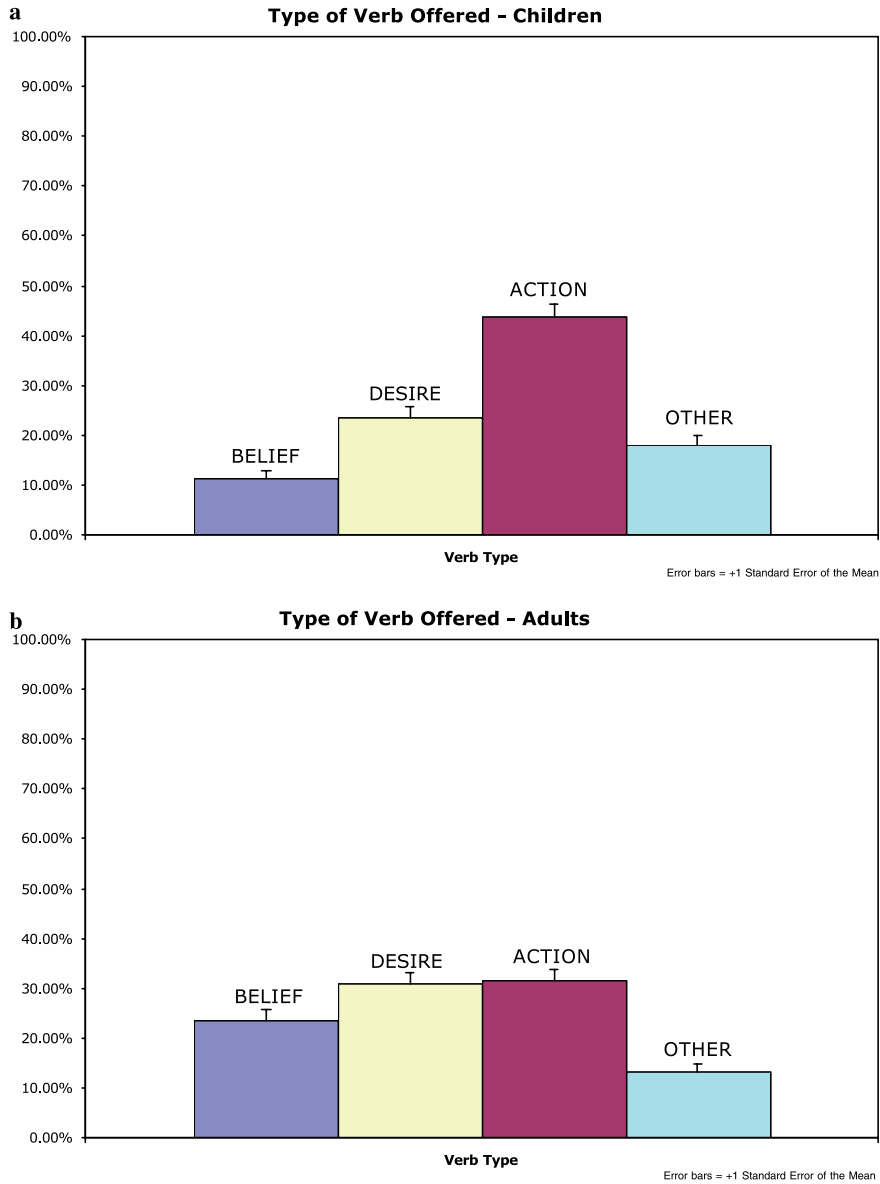
a

Type of Verb Offered – Children

Error bars = +1 Standard Error of the Mean

b

Type of Verb Offered – Adults

Error bars = +1 Standard Error of the Mean

Fig. 1. (a) and (b). Overall Percent of Verbs Offered in Exp. 3.

2.3.6.2. Going mental. Using the number of B verbs offered by children and adults as the dependent variable, we conducted two separate 2×2 repeated measures ANOVAs with Scene (TB, FB) and Syntax (S-comp, T) as within-subject variables. Results are presented in Figs. 2a and b (see also Tables 4a and 4b for a

a

% Belief Verbs Offered by Scene and Syntax – Children

Scene
- True Belief
- False Belief

Error bars = +1 Standard Error of the Mean

b

% Belief Verbs Offered by Scene and Syntax – Adults

Scene
- True Belief
- False Belief
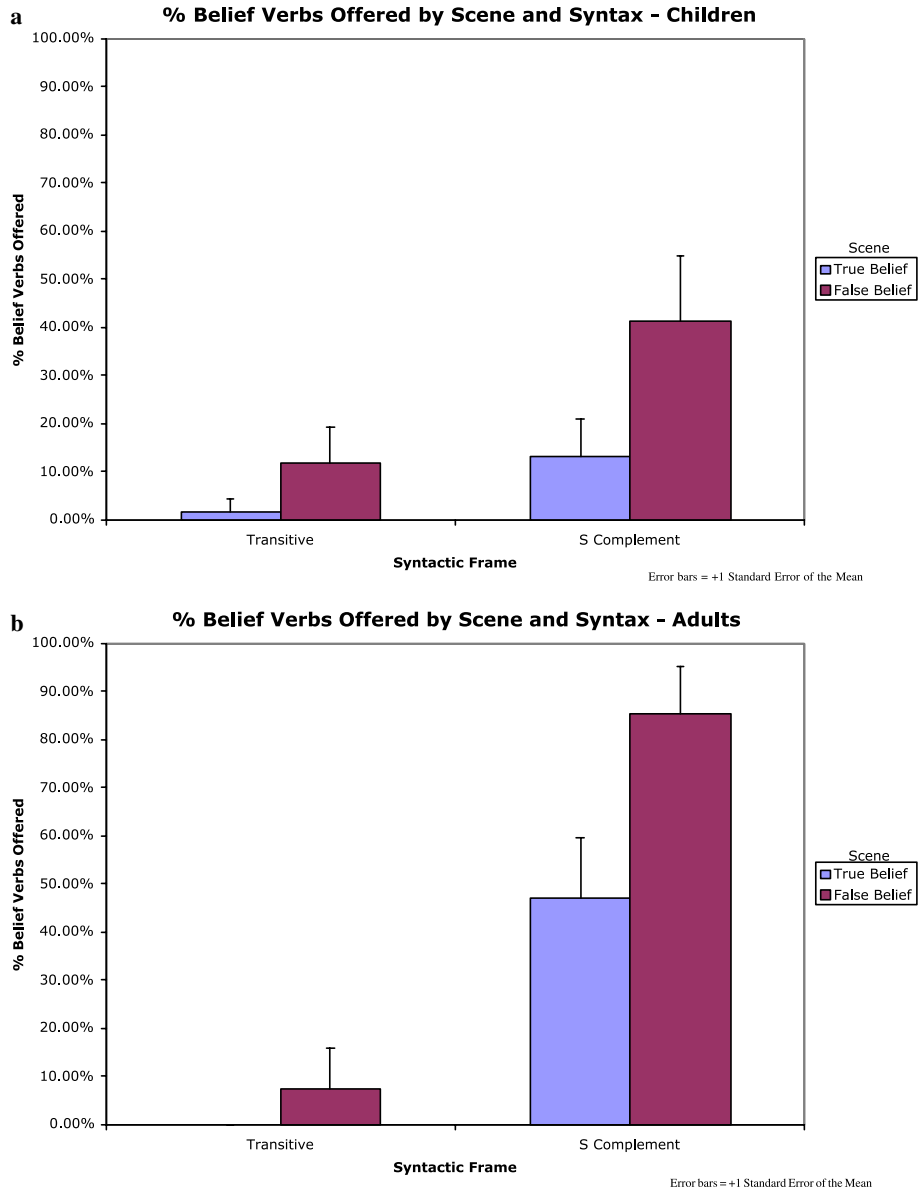
Error bars = +1 Standard Error of the Mean

Fig. 2. (a) and (b). Percent of Belief Verbs, by Scene and Syntax (Exp.3).

detailed breakdown of all verb responses). In both cases, the analyses revealed a significant main effect of Scene (children: $F(1, 33) = 21.78$, $p < .0001$; adults: ($F(1, 33) = 22.89$, $p < .001$)). Despite the strong bias toward action verb expression, attention to knowledge state (that is, the proportion of B responses) emerges from

Table 4a
Percent verb type by Scene and Syntax (Exp. 3) – children

	Belief	Action	Desire
Verb Type Offered – children			
Scene			
True Belief	7.4	58.8	3.7
False Belief	26.5	39.7	8.1
Desire	0	59	33.1
Syntax			
S Comp	27.2	29.4	6
Transitive	6.60	69	6
Infinitival	0	33	58.8
Scene and Syntax			
FB + S Comp	41.2	20.6	8.8
TB + S Comp	13.2	38.2	2.9
FB + Transitive	11.8	58.8	7.4
TB + Transitive	1.5	79.4	4.4
Desire + Infinitive	0	33.1	58.8

Table 4b
Percent verb type by Scene and Syntax (Exp.3) – adults

	Belief	Action	Desire
Verb Type Offered – adults			
Scene			
True Belief	23.5	49.3	1.5
False Belief	46.3	33.8	4.4
Desire	0.7	11.8	86.8
Syntax			
S Comp	66.2	16.9	2.9
Transitive	3.7	66.2	2.9
Infinitival	0.7	11.8	86.8
Scene and Syntax			
FB + S Comp	85.3	5.9	5.9
TB + S Comp	47.1	27.9	0.0
FB + Transitive	7.4	61.8	2.9
TB + Transitive	0.0	70.6	2.9
Desire + Infinitive	0.7	11.8	86.6

the background when the scene is one of false belief. For the adult subjects, these proportions almost double, rising from 23.5% in the TB situation to 46.3% in the FB situation. For the child subjects, the results are even more dramatic not only because the base rate of B verb response is so low for the TB situation (7.4%) but because it more than triples in the FB scenarios (26.5%). No wonder that mistaken states of knowledge and belief are the essence of theatre, both in farce and tragedy.

The ANOVAs also resulted in a significant main effect of Syntax (adults: $F(1,33) = 157.58$, $p < .001$; children: $F(1,33) = 24.59$, $p < .0001$). In scenes with a transitive prompt, B verbs occurred in only 3.7% of adult responses and 6.6% of the child responses. When presented with complement clause syntax, however, the number of B verbs increased in both adults and children (to 66.2% and 27.2% of all responses, respectively).

Finally, the analyses revealed an interaction between Scene and Syntax (adults: $F(1,33) = 15.27$, $p < .001$; children: $F(1,33) = 5.42$, $p = .026$). As Fig. 2 shows, the effects of cue combination lead differentially toward mental predicate conjectures. In detail, a transitive verb presented in the TB scene context never elicits a B verb by either adult or child subjects (0% and 1.5%, respectively). When the syntactic probe is S-comp, even in the TB situation, the rate of B verb responding rises very dramatically for adults (to 47.1%) but only slightly for children (to 13.2%). But when the listener receives both kinds of cue, the success in conjecturing a B verb interpretation is remarkably enhanced. For adults the effect is almost categorical when the situation is FB and the probe sentence is S-comp (85.5%). For the children the result is by no means as powerful absolutely, but the incidence of B verb conjectures triples compared to the TB situation (from 13.2% to 41%). Paired sample *t*-tests confirmed that the difference in percent of B verbs between the most helpful (FB scenes/S-comp) and least helpful (TB scenes/T frame) syntax-scene combinations is significant for both age groups (adults: $t(33) = -17.18$, $p < .001$; children: $t(33) = -3.95$, $p < .001$). Thus the combination of syntactic and situational evidence is compelling enough to support the identification of belief verbs, even in our young preschool participants. A comparison of B verbs in the two 'mixed' cue combinations (TB scenes with S-comp syntax vs. FB scenes with T syntax) yielded a significant difference for adults ($t(33) = 4.73$, $p < .001$) but not for children.[10]

An omnibus repeated-measures ANOVA on the number of B responses was conducted with Scene (TB, FB) and Syntax (S-comp, T) as within-subject factors and Age as a between-subject factor. Scene ($F(1,66) = 44.5$, $p < .001$), Syntax ($F(1,66) = 164.28$, $p < .001$), Scene × Syntax ($F(1,66) = 19.64$, $p < .001$), and Syntax × Age ($F(1,66) = 41.8$, $p < .001$) were all found to be significant. No other significant interactions were found. The Syntax × Age interaction was due to the higher impact of syntactic frame on adults' than on children's choices: as we saw earlier, B choices increased much more dramatically between the T- and S-comp

---

[10] Occasionally, children's B responses did not conform to strict frame compliance (e.g., *trick* was offered in response to a S-comp prompt even though the verb does not take a *that*-complement; see also analysis in next section). We thus conducted an additional repeated-measures ANOVA over B responses after strictly non-frame-compliant responses had been excluded from the child data. The results did not change: we found significant main effects of Scene and Syntax ($F(1,33) = 6.23$, $p = .018$; $F(1,33) = 30.5$, $p < .001$, respectively), and a significant interaction ($F(1,33) = 5.58$, $p = .024$).

conditions in adults (from 3.7% to 66.2%) than in children (from 6.6% to 27.2%).[11]

2.3.6.3. Saying it right. The percentage of strictly frame compliant responses offered by adults and children is given in Figs. 3a and b. An omnibus analysis using the number of such responses as the dependent variable and Scene, Syntax and Age as factors resulted in significant main effects of Scene ($F(1, 66) = 4.64$, $p = .035$) and Syntax ($F(1, 66) = 70.35$, $p < .001$), as well as interaction effects of Scene × Syntax ($F(1, 66) = 8.32$, $p = .005$) and Syntax × Age ($F(1, 66) = 65.58$, $p < .001$). A Scene × Age interaction just missed significance ($F(1, 66) = 3.4$, $p = .069$).

We conducted two separate repeated measures ANOVAs over the number of strictly frame compliant responses offered by adults and children with Scene and Syntax as factors. Strict frame compliance in adults was nearly at ceiling, so the analysis resulted in no significant main or interaction effects. For children, the analysis yielded significant main effects of Scene ($F(1, 33) = 4.55$, $p = .04$) and Syntax ($F(1, 33) = 84.22$, $p < .001$), as well as a significant interaction of Scene and Syntax ($F(1, 33) = 5.69$, $p = .023$). The effect of Scene was due to a greater level of strict frame compliance for FB than for TB scenes (53.7% vs. 44.1%, respectively). As for the effect of Syntax, S-comp frames met with strict frame compliance only 27.9% of the time, while T frames had a strict frame compliance level of 69.9%. The Scene × Syntax interaction arose because children were likely to be strictly frame-compliant for T frames regardless of the scene properties (70.6% for TB vs. 69.1% for FB scenes); however, for S-comp frames, there was a significant difference between TB and FB scenes (17.6% vs. 38.2%; $t(33) = 3.23$, $p = .003$).

2.3.7. Discussion

Results of this experiment reveal the importance of both observational and syntactic cues for the identification of mental-state predicates. As in Experiment 1, FB scenes enhance the salience of the actors' inferred states of mind, as measured by a significantly higher proportion of mental-state interp7retations of novel verbs compared to TB scenes (26.5% vs. 7.4% in children's responses and 46.3% vs. 23.5% in adults' responses). Syntactic information places even more striking constraints on the

---

[11] A pair of repeated-measures ANOVAs over action verb responses offered by the two subject groups with Scene (TB, FB) and Syntax (S-comp, T) as factors resulted in a significant main effect of Scene (adults: $F(1, 33) = 9.72$, $p < .001$; children: $F(1, 33) = 21.78$, $p < .001$) and Syntax (adults: $F(1, 33) = 117.85$, $p < .001$; children: $F(1, 33) = 70.34$, $p < .001$), but no significant interaction (see Table 4). Both groups offered action verbs more frequently in TB than in FB scenes (adults: 49.3% vs. 33.8%; children: 58.8% vs. 39.7%) and in T frame prompts than in S-comp prompts (adults: 66.2% vs. 16.9%; children: 69% vs. 29.4%). Also both groups preferred A verb hypotheses for TB scenes with T frames (adults: 70.6%; children: 79.4%) and FB scenes with T frames (adults: 61.8%; children: 58.8%). This preference decreases when TB scenes appear with B-biasing (S-comp) syntax (adults: 27.9%; children: 38.2%) and especially when both observational cues (FB scenes) and syntax (S-comp) supported B guesses (adults: 5.9%; children 20.6%). Despite their variability, these results show that syntax trumps scenes in interpreting novel verbs for both adults and children. An omnibus ANOVA adding Age as a factor revealed only main effects of Scene ($F(1, 66) = 44.5$, $p < .0001$) and Syntax ($F(1, 66) = 184.05$, $p < .0001$) and no interactions.
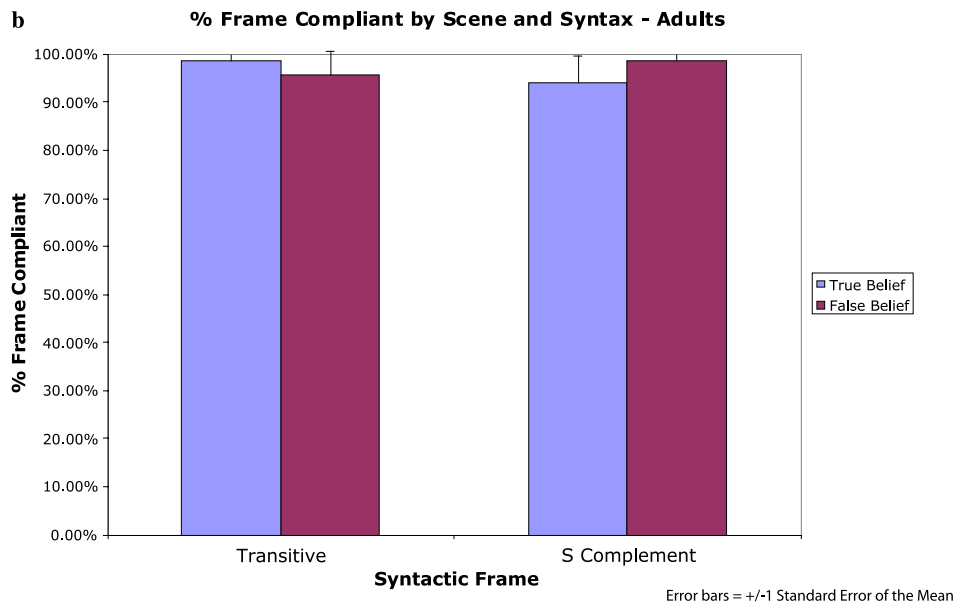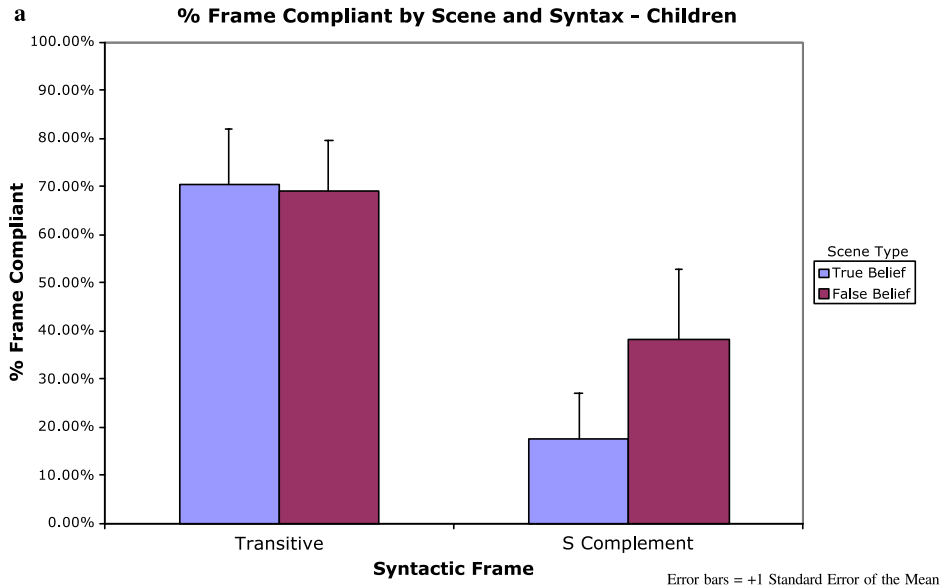
a

b

Fig. 3. (a) and (b). Percent of strict frame compliance, by Scene and Syntax (Exp. 3).

preferred construal of a scene: put differently, syntax acts as an attentional zoom lens
(Gleitman, 1990) by bringing different interpretations into the listener's mental focus.
While transitive frames with direct object-NPs almost never occurred with belief
verbs, *that*-complements strongly encouraged belief verb construals (27.2% and

66.2% of all responses in children and adults, respectively). In trials where observational and syntactic cues converged, nearly half (41.2%) of children's responses and the overwhelming majority (85.5%) of adults' responses were belief verbs.

This experiment allows us to compare the use of extra-linguistic and linguistic-structural cues in adults and children in ways that we think are suggestive about the mapping problem as solved *in vivo*. A possible outcome was for adult responses to be more heavily influenced by the linguistic input and for children's responses to be determined more by the nature of the scenarios themselves. Such a finding would tend to support the position that mental state verbs are learned relatively late because the maturation (or construction) of the relevant concept space stretches over a long period and is not available for this task in younger verb learners. However the finding is that adults and children were sensitive to the same variables in the same approximate difference-magnitudes, lending some *prima facie* support to the information-structure interpretation of verb learning. Participants in both age groups converge on similar interpretations when given the same combination of extra-linguistic and linguistic cues (Figs. 2a and b). This is not surprising in light of our analysis of the mapping problem. Old or young, our subjects are all novices in respect to gleaning the meaning of *gorp* and *glick*. If the procedures for solving these mapping problems for the vocabulary implicate the recruitment of both syntactic and situational cues, the findings in our experimental model are much as would be expected.

At the same time there are detailed differences in the adult and child response styles that are worth noting. First, child subjects provide a more variable picture than adults. To be sure, the rate of word learning in the real world is somewhat higher for older children and adults than it is for the preschool population we tested (Bloom, 2000) but it is likely that the rather large difference we see in our data is more indicative of adults' better understanding of laboratory task demands than a difference related to real word-learning rates. More likely to reflect the real facts of the matter is the greater potency of sentence complementation for eliciting mental-content responses from adults than for children (Table 4). Though all the participants are sensitive to this kind of language-internal cue, its relative weight is greater for the adults than for the children. Because the syntactic privileges of mental verbs are an important source of evidence for their meaning, such difficulties with complementation structures can explain the delay in the acquisition of mental verbs.

3. General discussion

Ryan (age 3 years) returns from the pizza counter carrying a soda cup.

Mom: *Put that down. You are not allowed to have soda. Put that down.*
Ryan: *Why?*
Mom: *You are not allowed to have soda. Put it down.*
Ryan: *It's not soda. The pizza man gave me water.*
Mom: *Oh….I'm sorry. I thought it was Coke. See…. there is a Coke sign outside, so I thought it was Coke.*
Ryan: *Oh that was funny. You were mad. Let's do that again.*

3.1. Summary and discussion of the findings

This article posed the problem of why some words are harder for young children to acquire than others, even when they are frequent in the speech of caregivers. We chose the mental-content verbs to investigate this problem, just because they are iconic instances of such "hard" words. Two approaches were considered as accounts of the special difficulty of these words. The first was that preschoolers lack the conceptual wherewithal to think very well about thinking itself, with the consequence that they cannot acquire words that refer to mental predicates. The second posited that even if young children are adequate in this regard, they – and in fact, all the word learners, young and old – should find mentalistic terms difficult as a consequence of the initial procedure for solving the mapping problem for vocabulary; namely, by inspecting the extralinguistic contexts for a word's utterance (word-to-world pairing). Mental content words are not efficiently identified using this procedure because the situational contexts for talking about states and acts of mind are variable and subtle, when available to inspection at all.

Of course *both* linguistic and conceptual issues must play roles in understanding just when and how different aspects of the lexical stock of a language are acquired, so it is not our explanatory aim to choose between them; rather, our goal is to see how solutions to the word learning problem are apportioned between these two factors. As a first means of disentangling their contributions (and interactions), we employed adaptations of the so-called Human Simulation Paradigm. This kind of manipulation is designed to unconfound the conceptual versus informational factors by studying adults (so as to remove conceptual inadequacy from the learning equation) while varying their informational circumstances (so as to study the influence of increasing linguistic expertise on the procedures for vocabulary acquisition). Our first task was to discover extralinguistic circumstances that would naturally elicit the use of mental-content terms, especially from caregivers to their young offspring. We found our inspiration in that regard in an interchange between one of us (KC) and her 3-year-old son, which appears as the epigram to this discussion ("I *thought* it was Coke. See?…"). Situations of false belief seem to be quintessential triggers for talk of mental content.[12]

Experiment 1 investigated this hypothesis by asking adults "What might you say?" when confronted with videotapes depicting characters in both true belief and false belief states, compared with scenarios that are evidently most easily interpreted as depicting simple actions (e.g., drinking) or states of attempted desire-satisfaction (e.g., trying). As we showed (see Table 1), this conjectured talk overwhelmingly centers on the event – the action: almost no matter the content of that event, our

---

[12] Supportive evidence comes from a picture-book reading experiment (Ratner & Olver, 1998). These investigators recorded mothers' spontaneous conversations (interruptions of the story reading) with their 4-year-old children as they read a folk tale that involved deception. Approximately 80% of the utterances where parents made reference to thoughts and feelings throughout the entire book reading occurred during the interruptions where the mothers discussed the deceptions of the story.

participants uttered verbs expressing concrete actions over 90% of the time, and for all scenario types. All the same, scenarios that emphasized the character's state of belief (true or false) also or alternatively elicited mental predicates about 30% of the time overall, and 42% of the time in the scenarios of false belief. Interestingly enough, exaggerated situations of pratfalls and risible confusions were neither required nor even especially useful in this regard, and so the "zaniness" factor that we probed as a way of upping the salience of false belief situations had no added effect. That there is a discernible false belief, even a quite prosaic one, seems to be sufficient.

Insofar as this first simulation is suggestive of relevance conditions for mental verb usage, we conclude:

(i) Actions trump beliefs in the interpretation of the extralinguistic world.
(ii) Observations that suggest the presence of a false belief can mitigate this bias and thus trigger talk of belief from adults.

Experiment 2 investigated the conditions triggering likely (adult) utterances of mental predicates in further ways. The experimental question now concerned the informational conditions that give rise to mental predicate conjectures for novel words. So the second innovation here was to offer subjects either (a) extralinguistic information of a kind suited to belief predicates (false belief) or not so suited (action scenarios); or (b) linguistic information of a kind suited to belief predicates (S-comp constructions) or not so suited (transitive constructions), or (c) both (a) and (b) together.

More precisely, we crossed situation (Action or FB) and syntax (S-comp or Transitive) to see the effects of these information variables when they were present singly or jointly (either pitted against each other, or conspiring). So as not to contaminate the results with concordant distributional information, in the linguistic conditions we substituted nonce words for nouns as well as for the "mystery verb."

The findings from this study strongly support the view that mental-predicate conjectures are insalient as interpretations of the passing scene. Any excuse for an action interpretation was seized upon by our participants. This excuse could be either (or both) an action scene or a transitive construction. When any of these informational components were available, the proportion of belief verb conjectures was very low and about the same whatever the accompanying information (between 3% and 9%). A moment's reflection makes clear why this should be so. Every situation depicting a human (well, at least most humans, and if not comatose) implicates some state or states of mind, lots of them probably quite irrelevant to the present action. So it takes some special feature or property of that situation to suggest that these invisible mental states are somehow, in some special case, central to the ongoing discourse. The false belief situation seems to fit this bill. In the experimental false belief scenes, the proportion of belief verb responses increased (from 6.5% to 40.33% overall).

Special linguistic-descriptive properties were found to be even more effective in increasing the likelihood of a mentalistic conjecture. Belief verb response proportion increased most in the S-comp syntax condition (to 55%), even without any supportive

(or, for that matter, conflicting) evidence from an observed situation. Finally, false belief scenarios paired with S-comp structures resulted in belief-verb conjectures roughly 80% of the time. Considering the fact that these participants had only singulary information about either kind of information source – they saw only one situation and/or only one syntactic frame and so could not do cross-situational or cross-constructional inferencing – the power of these information sources in this simulation is quite surprising. Our conclusions from this second experiment are:

(iii) Mental verb syntax (S-comp) elicits mental verb interpretation of novel words reliably more often than even the strongest (false-belief) situational cue.
(iv) Combining these two cue types is just about decisive for a mental predicate interpretation.

Experiment 3 extended these findings to children in the first major stage of predicate-term acquisition. There are three major findings. First, the priority of action over mentalistic interpretation, both for the children and their adult controls, shines through in every analysis. Overall both populations offer more action conjectures by far (Fig. 1); for both populations this priority of action responses is weakened in the presence of false belief situations (Table 4) and belief constructions (Table 5). And the conspiracy of these two kinds of cue has a very strong, overadditive effect in eliciting mentalistic conjectures (Fig. 2). All these similarities across populations suggest that the character of vocabulary acquisition is materially explained as a learning problem with its own internal logic, one that is responsive to accumulating evidence and to re-representation of the problem space via the enrichment of evidentiary resources.

At the same time, we may also be seeing the footprint of cognitive-maturational factors in these results. Comparison of Figs. 2a and b shows that children are only half as reliable as adults in offering mental-state conjectures when these would be appropriate. False belief story lines as well as belief syntax (S-comp) are both less effective for the children than for the adults in eliciting the mentalistic response. Part of this difference can be written off as general population differences in test-taking ability and laboratory-task experience. But not all of it: the child subjects, unlike adults, are more story-compliant if the story is an action type story than if it is a false-belief type story (Table 4); and the child subjects, unlike adults, are more frame-compliant if the syntax is of the action-interpretation sort (that is, transitive) than if it is of the mental-interpretation sort (Fig. 3). This suggests a disinclination to focus attention on mental states that, as many investigators of theory of mind have suggested, may be a cognitive hallmark of early childhood.

We have interpreted our findings as suggesting a learning procedure for vocabulary that has three crucial properties: it is sensitive to different types of information in hypothesizing the meaning of novel words; it is responsive to the presence of multiple conspiring cues; it especially weights the language-internal cues when faced with unreliable extralinguistic cues. These properties heavily influence response style for novel words in both adult and child learners. The results thus bolster a position we have also urged elsewhere, to the effect that the character of novice vocabulary is

most naturally explained as having more to do with inherent properties of a learning procedure than with significant early gaps in the conceptual repertoire (Gleitman et al., 2005). From this perspective, young children are just a special case of novice learners. Their delay in acquiring certain kinds of term is attributable to ignorance of the distinctive linguistic devices that bring these into attentional focus.

Yet obviously there are limits to the scope of a theory of word learning whose only parameters are learning-theoretic. If we had expanded our stimulus set to theory-driven concepts and expressions like *prime number* or *plutocracy*, no doubt differences in the sheer representational and reasoning skills of adults versus children would have come into play and borne more of the explanatory burden. Even for the basic mental terms studied here we have documented discernible age differences in the kind of category that can readily be brought to the forefront of attention. In every manipulation, mental conjectures were globally less accessible to the preschoolers than to the adults. It is not a matter of not "having" such concepts but only a matter of their lesser availability and prominence, as we saw by the fact that stronger environmental and linguistic cues can elicit such responses from the preschool participants.

Speaking now more generally, such effects of salience – rather than concept gaps – are seen again and again in studies of which words are learned, when they are learned, and the conditions for their learning. Consider as another kind of example the case of color perception and color terms. By the age of 4 months infants have extensive experience with colors, discriminate them well, and also show preferences among them (Bornstein, 1978 *inter alia*). Yet as Carey (1982) pointed out, even though color is undoubtedly a perceptual primitive, a term like *red* rarely comes up in infants' minds as a possible candidate for lexicalization and so *red* is usually learned later than terms even for adventitious object properties such as *clean* and *dirty* (Bartlett, 1977). This insalience of color terminology for young children was perhaps first remarked upon by Charles Darwin:

> I carefully followed the mental development of my small children, and I was astonished to observe in two or, as I think, three of these children, soon after they had reached the age in which they knew the names of all the ordinary things, that they appeared to be entirely incapable of giving the right names to the colors of a color etching. They could not name the colors, although I tried repeatedly to teach them…I remember quite clearly to have stated that they are color blind (Darwin, 1877, pp. 367–76) .

Rice (1980) attempted to teach the terms *red* and *green* to 3-year-olds who did not know any color terms. She reported that it took from 500 to 2000 trials distributed over three weeks to accomplish this. It is as if children are astounded at the very idea that a language would have terms for the hues of things. Many related results appear in the literature of vocabulary acquisition: young children take shape as a more salient cue than size or color (Landau, Smith, & Jones, 1998), they favor wholes over parts (Markman & Wachtel, 1988), objects over their component properties (Hall, Waxman, & Hurwitz, 1993), and so forth. In none of these cases can it be argued that the children do not "see in color," "discriminate sizes," etc. Rather, the word-learning

effects for such categories seem to be a matter of differential salience, sometimes differential salience that pertains only or primarily in the context of word learning; that is, when the child's task is to match some world feature with a novel linguistic label (Imai & Gentner, 1997; Markman & Hutchinson, 1984). In these cases too, much as with the mental terms we have studied here, learning is successful when distinctive linguistic properties trigger attention to specific category types (for the case of adjectives, see, e.g., Waxman & Klibanoff, 2000; Mintz & Gleitman, 2002).

3.2. Final thoughts

Developmental psychologists often assume, as have we, that children embark upon the task of word learning with the advantage of being "creatures like us." This gives them practical (though doubtless nonconscious) ways to appreciate the categories, motives, and discourse intentions of their caregivers so as to understand what's being said to them. Armed with these mind-reading capacities and inclinations, infants can acquire word meanings despite the buzzing blooming confusion of things, events, properties, and relations that confront them in the environment (see e.g., Baldwin, 1991; Bloom, 2000). Other commentators assume, quite to the contrary, that at the beginning of the word learning process children are cognitively very different from adults: that among other things, they are unable to reason about beliefs and other mental contents, and hence cannot even attribute the correct referents to mental state expressions (e.g., Perner, 1991). Sometimes, both of these views are held at the same time: children are supposed to be solving the conceptual puzzles of mental concepts around the same time as they are figuring out the meanings for epistemic expressions. These opposed perspectives on the learning problem for vocabulary are a long way from final resolution. Here we have tried to offer some directions by focusing on the special case of mental terms. Taken at their broadest our results suggest considerable continuity of conceptual life and learning procedures as between adults and children. But to learn a word, to find out what *gorp* and *pilk* label in the ambient world, one has to jog the mind, to focus one's attention on a particular perspective or description among the many that are always available. Looking out at the world we observe entities and their actions, not beliefs, unless these latter are patently absurd. Listening, we hear descriptions of entities and their acts, unless these descriptions patently mismatch the semantics of syntactic structure. It is primarily under these unusual psychological and syntactic circumstances that we think about thinking itself.

Acknowledgments

Appendix 1. Verb types offered in Exp. 1

Belief Verbs: acknowledge, be absorbed in/aware/unaware/bewildered/careful/creative/interested/prepared/preoccupied, be/make sure, bother, concentrate, confuse, decide, distract, engross, expect, fantasize, figure out, focus, forget, get/be absent minded, guess, ignore, imagine, know, learn, make a mistake/mistake, mean, notice, overestimate, pay attention, play a joke, play a prank/pull a prank, pretend, realize, remember, remind, represent, stump, surprise, think, trick/play a trick, understand, wonder, worry, zone out

Desire Verbs: attempt, be afraid/angry/ashamed/bored/disgusted/frustrated/happy/impulsive/patient/respectful, enjoy, entertain, get angry/mad/frustrated, hate, like, look for, need, respect, trust, try, want, wish

Action Verbs: act, aid, answer, arrive, ask, be organized, begin, belabor, bet, bounce, break, bring, browse, burn, buy, call, catch, cause, change, check, choose, clean, close, color, come, complain, complete, consume, contain, continue, control, cook, correct, create, cross out, design, die, disappear, disturb, divide, draw, drink, dress, drop, dump, dust, eat, enable, end, enter, erase, examine, exchange, expand, experiment, explain, express, facilitate, fail, fall, fend, fill, find, finish, finger-paint, fit, fix, flip, float, fly, follow, force, gather, get (ready/stuck), give, give up, go, grab, grow, hang, harm, heat, help, hide, highlight, hit, hurt, infringe, interrupt, jam, jump, keep, kick, kill, knock, knock over, land, lay, leave, lift, listen, live, look, lose, make, manage, mess, mess around/up, microwave, misplace, miss, mix, move, offer, open, organize, overlook, paint, participate, pass, pause, perform, persevere, phone, pick, place, play, poison, pour, practice, prepare, problem solve, prove, pull, punish, push, put, reach, react, read, rearrange, recap, relax, reorganize, replace, resume, retrieve, return, review, ride, ring, risk, roll, run, run out of, rush, sabotage, say, scratch, scribble, search, see, select, separate, sip, sit, skim, slap, slide, slip, smile, sneak, soar, solve, sort, spend, spill, spin, spit, spread, stand, stare, start, stay, steal, step, stick, stop, store, straighten, stretch, stress, struggle, study, stumble, substitute, succeed, suggest, swallow, switch, take, take care, talk, taste, tell, test, thank, throw, tidy, toss, touch, trip, turn (into), type, undress, use, visit, wait, wake up, walk, warm (up), watch, waste, wear, win, wipe, work, write

Other Verbs: allow, appear, be, belong, do, get along, happen, have (got), keep, let, make, matter, seem, serve (one right), settle

Appendix 2a. Verb types offered in Exp. 3 – children

Verb types are listed along with number of occurrences in participants' responses and in a corpus of 1,014,232 words of English (Kučera & Francis, 1967).

Verb category‘			
Action	Belief	Desire	Other
Ask (3) 128	Be surprised (1) 58	Like (2) 1290	Be (54) 6377
Bang (1) 7	Dream (2) 64	Love (8) 232	Do (7) 1363

(*continued on next page*)

Appendix 2 (continued)

Verb category			
Action	Belief	Desire	Other
Beg (1) 11	Fall for (1)	Need (2) 360	Feel better (1)
Break (18) 88	Guess (1) 56	Try (8) 140	Have to (1)
Buy (1) 70	Know (3) 638	Want (75) 329	Make (10) 794
Clean (2) 70	Lie (2) 59	Wish (1) 110	Might be (1)
Climb (1) 2	Pretend (2) 8		
Cry (4) 48	Think (28) 433		
Cut (2) 192	Trick (6) 15		
Drop (19) 59			
Eat (39) 61			
Fall (3) 147			
Get (7) 750			
Give (9) 391			
Go (10) 626			
Grow (1) 63			
Help (1) 311			
Hold (1) 169			
Hug (10) 3			
Hurt (1) 37			
Jump (1) 24			
Leave (3) 205			
Lift (3) 23			
Look (2) 399			
Make (3) 794			
Open (2) 319			
Peel (3) 3			
Pick (3) 55			
Pound (1) 28			
Pour (1) 9			
Reach (10) 106			
Rest (1) 163			
Say (4) 504			
Sleep (1) 65			
Split (1) 30			
Take (6) 611			
Talk (1) 154			
Thank (1) 36			
Throw (1) 42			
Tuck (1) 2			
Turn into (1)			
Unlock (1) 3			
Water (3) 442			
Wear (1) 36			

Total: 408 items
   (includes 17 null/no verb responses)

Appendix 2b. Verb types offered in Exp. 3 – adults

Verb category			
Action	Belief	Desire	Other
Apologize (1) 1	Agree (1) 51	Desire (4) 79	Be (22) 6377
Beg (1) 11	Appreciate (3) 26	Hope (1) 178	Have (30) 3941
Bite (4) 10	Apprehend (1) 1	Like (7) 1290	Make (2) 794
Break (9) 88	Assume (1) 63	Long (1) 755	
Bring (5) 158	Be fooled (4) 3	Love (15) 232	
Bruise (1) 3	Be surprised (2) 58	Miss (1) 258	
Cry (2) 192	Believe (13) 200	Need (1) 360	
Deliver (1) 18	Dream (2) 64	Try (7) 140	
Drop (15) 59	Know (4) 638	Want (73) 329	
Eat (6) 61	Make a conjecture (1)	Wish (17) 110	
Find (1) 399	Notice (3) 59		
Gasp (1) 3	Realize (2) 69		
Give (8) 391	Think (62) 433		
Give up (1)			
Grin (1) 13			
Hold (4) 169			
Hug (2) 3			
Leave (2) 205			
Lift (8) 23			
Look (3) 399			
Offer (2) 80			
Peel (7) 3			
Reach (14) 106			
Say (4) 504			
See (15) 722			
State (1) 808			
Stretch (1) 26			
Stub (1) 3			
Take (8) 611			
Total: 408 items (includes 3 null responses)			

References

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264.

Au, T., Dapretto, M., & Song, Y.-K. (1994). Input versus constraints: early word acquisition in Korean and English. *Journal of Memory and Language, 33*, 567–582.

Baker, M. C. (2001). *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

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

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

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–609.

Bartlett, E. (1977). The acquisition of the meaning of color terms: A study of lexical development. In P. Smith & R. Campbell (Eds.), *Recent advances in the psychology of language, IV* (pp. 89–108). New York: Plenum Press.

Bartsch, R., & Wellman, H. (1995). *Children talk about the mind*. Oxford: Oxford University Press.

Bloom, L., Lightblown, P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development, 40*, 1–97.

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

Bloom, P. (2002). Mindreading, communication, and the learning of the names for things. *Mind and Language, 17*, 37–54.

Bloom, P., & German, T. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, B25–B31.

Bornstein, M. H. (1978). Chromatic vision in infancy. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 12, pp. 117–182). New York: Academic Press.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: the acquisition of an explicit theory of mind. *Developmental Psychology, 18*, 906–921.

Brown, P., & Dell, G. (1987). Adapting production to comprehension: the explicit mention of instruments. *Cognitive Psychology, 19*, 441–472.

Brown, R. (1957). Linguistic determinism and the parts of speech. *Journal of Abnormal and Social Psychology, 55*, 1–5.

Bruner, J. (1974). The ontogenesis of speech acts. *Journal of Child Language, 2*, 1–19.

Carey, S. (1982). Semantic development: The state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art*. New York: Cambridge University Press.

Carruthers, P., & Smith, P. (Eds.). (1996). *Theories of theories of mind*. Cambridge: Cambridge University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Holland: Foris.

Croft, W. (1990). *Typology and universals*. New York: Cambridge University Press.

Darwin, C. (1877). A biographical sketch of a young child. *Kosmos*, 1, 367–376. As cited in Bornstein, M. H., Chromatic vision in infancy. In H. Reese & L. Lipsitt (Eds.), *Advances in child development and behavior*, Vol. 12. NY: Academic Press.

de Villiers, J., & de Villiers, P. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind*. Hove, UK: Psychology Press.

Dromi, E. (1987). *Early lexical development*. Cambridge: Cambridge University Press.

Fillmore, C. J. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). New York: Holt, Rinehart and Winston.

Fisher, C. (1996). Structural limits on verb mapping: the role of analogy in children's interpretation of sentences. *Cognitive Psychology, 31*, 41–81.

Fisher, C., & Gleitman, L. R. (2002). Breaking the linguistic code: Current issues in early language learning. In C. R. Gallistel (Ed.), *Learning, motivation, and emotion* (Vol. 3). *Stevens' handbook of experimental psychology*. New York: John Wiley & Sons.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology, 23*, 331–392.

Fisher, C., Hall, G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: syntactic and conceptual constraints on vocabulary growth. *Lingua, 92*, 333–375.

Gentner, D. (1978). On relational meaning: the acquisition of verb meaning. *Child Development, 49*, 988–998.

Geyer, H. L. (1994). Subcategorization as a predictor of verb meaning: evidence from Hebrew. Unpublished manuscript, University of Pennsylvania.

Gillette, J., Gleitman, H., Gleitman, L. R., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*, 135–176.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 3–55.

Gleitman, L. R. (2003). For Zellig Harris: verbs of a feather flock together II. In B. Nevin, H. Hoenigswald, & H. Hiz (Eds.), *The legacy of Zellig Harris*. Amsterdam: Benjamins.

Gleitman, L., & Newport, E. (1996). Universals of language learning. In L. Gleitman & M. Liberman (Eds.), *Language* (Vol. I). *Invitation to cognitive science*. Cambridge, MA: MIT Press.

Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. (2005). Hard words. *Language Learning and Development, 1*, 23–64.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Goldberg, A. (2004). But do we need universal grammar? Comment on Lidz et al. (2003).. *Cognition, 94*, 77–84.

Goldin-Meadow, S. (2003). The resilience of language: What gesture creation in deaf children can tell us about how all children learn language. In J. Werker & H. Wellman (Eds.), *Essays in developmental psychology*. New York: Psychology Press.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts and theories*. Cambridge, MA: MIT Press.

Grice, P. (1967). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics: Vol. 3. Speech acts* (pp. 41–58). New York: Academic Press.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., & Wilson, R. (1989). The learnability and acquisition of the dative alternation in English. *Language, 65*(2), 203–257.

Gruber, J. S. (1965). Look and see. *Language, 43*(4), 937–947.

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: a training study. *Developmental Science, 6*, 346–359.

Hall, D. G., Waxman, S. R., & Hurwitz, W. M. (1993). How 2- and 4-year-old children interpret adjectives and count nouns. *Child Development, 64*, 1661–1664.

Harris, Z. (1951). *Methods in structural linguistics*. Chicago: University of Chicago Press.

Hume, D. (1739–40). *A treatise of human nature*. Reprinted from Oxford University Press, 1978.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). The emergence of action categories in the child: evidence from verb meanings. *Psychological Review, 90*, 72–93.

Imai, M., & Gentner, D. (1997). A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62*, 169–200.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Johnson, C. N., & Maratsos, M. P. (1977). Early comprehension of mental verbs: think and know. *Child Development, 48*, 1743–1747.

Kučera, H., & Francis, N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Landau, B., Smith, L. B., & Jones, S. (1998). Object shape, object function, and object name. *Journal of Memorv and Language, 38*, 1–27.

Lederer, A., Gleitman, H., & Gleitman, L. (1995). Verbs of a feather flock together: Semantic information in the structure of maternal speech. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things* (pp. 277–297). Hillsdale, NJ: Erlbaum.

Leslie, A. (1994). Pretending and believing: issues in the theory of ToMM. *Cognition, 50*, 265–288.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences, 9*, 459–462.

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*, 45–85.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences, 8*, 528–533.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Li, P. (1994). Subcategorization as a predictor of verb meaning: Cross-linguistic study in Mandarin. Unpublished ms., University of Pennsylvania.

Lidz, J., & Gleitman, L. R. (2004). Yes, we still need universal grammar (Reply to Goldberg, 2004). *Cognition, 94*, 85–93.

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition, 87*, 151–178.

Locke, J. (1690). *An essay concerning human understanding* Repr. in 1964.. Cleveland: Meridian Books.

Markman, E., & Wachtel, G. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*, 121–157.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: taxonomic vs. thematic relations. *Cognitive Psychology, 16*, 1–27.

McCawley, J. (1968). The role of semantics in a grammar. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory* (pp. 124–169). New York: Holt, Rinehart & Winston.

Mintz, T., & Gleitman, L. R. (2002). Adjectives really do modify nouns. *Cognition, 84*, 267–293.

Moore, C., & Furrow, D. (1991). The development of the language of belief: The expression of relative certainty. In D. Frye & C. Moore (Eds.), *Children's theories of mind: Mental states and social understanding* (pp. 173–193). Hillsdale, NJ: Erlbaum.

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development, 60*, 167–171.

Naigles, L. (2000). Manipulating the input: Studies in mental verb acquisition. In B. Landau, J. Sabini, J. Jonides, & E. Newport (Eds.), *Perception, cognition and language* (pp. 245–274). Cambridge, MA: MIT Press.

Naigles, L. G. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*, 357–374.

Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: defining a role for syntax. *Child Development, 64*, 1665–1687.

Naigles, L., Gleitman, H., & Gleitman, L. R. (1992). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and cognition: A developmental perspective* (pp. 104–140). Norwood, NJ: Ablex.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.

Papafragou, A., Li, P., Choi, Y., & Han, C. (in press). Evidentiality in language and cognition. *Cognition*.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition, 98*, B75–B87.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Pinker, S. (1989). *Learnability and cognition*. Cambridge, MA: MIT Press.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515.

Ratner, N. K., & Olver, R. R. (1998). Reading a tale of deception, learning a theory of mind? *Early Childhood Research Quarterly, 13*, 219–239.

Resnick, P. (1996). Selectional constraints: an information-theoretic model and its computational realization. *Cognition, 61*, 127–159.

Rice, M. (1980). *Cognition to language: Categories, word meanings, and training*. Baltimore: University Park Press.

Senghas, A., Coppola, M., Newport, E. L., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development* (Vol. 2). Somerville, MA: Cascadilla Press.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental terms: A systematic investigation of the first reference to mental state. *Cognition, 14*, 201–321.

Smiley, P., & Huttenlocher, J. E. (1995). Conceptual development and the child's early words for events, objects, and persons. In M. Tomasello & W. Merriman (Eds.), *Beyond names for thing: Young children's acquisition of verbs* (pp. 21–62). Hillsdale, NJ: Erlbaum.

Smith, L. B. (2000). Avoiding association when it's behaviorism you really hate. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Breaking the word learning barrier* (pp. 169–174). Oxford: Oxford University Press.

Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts. In G. Hall & S. Waxman (Eds.), *Weaving a lexicon* (pp. 257–293). Cambridge, MA: MIT Press.

Snedeker, J. & Li, P. (2000). Can the situations in which words occur account for cross-linguistic variation in vocabulary composition? In J. Tai & Y. Chang (Eds.), *Proceedings of the Seventh International Symposium on Chinese Languages and Linguistics*.

Snedeker, J., Li, P., & Yuan, S. (2003). Cross-cultural differences in the input to early word learning, *Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Snedeker, J., Thorpe, K., & Trueswell, J. (2001). On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. In J. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society* (pp. 964–969). Mahwah, NJ: Erlbaum.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (2nd ed.). Harvard: Harvard University Press.

Tardiff, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns vs. verbs: a comparison of English, Italian and Mandarin. *Journal of Child Language, 24*, 535–565.

Tomasello, M., & Kruger, A. (1986). Joint attention on action: acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19*, 311–333.

Trueswell, J., & Gleitman, L. R. (2004). Children's eye movements during listening: evidence for a constraint-based theory of parsing and word learning. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Urmson, J. (1963). Parenthetical verbs. In C. Caton (Ed.), *Philosophy and ordinary language* (pp. 220–246). Urbana: University of Illinois Press.

Vendler, Z. (1972). *Res cogitans*. Ithaca, NY: Cornell University Press.

Waxman, S. R., & Lidz, J. (2006). Early word learning. In D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology: Cognition, Perception and Language* (6th ed., pp. 299–335). New York: Wiley.

Waxman, S. R. (1994). The development of an appreciation of specific linkages between linguistic and conceptual organization. *Lingua, 92*, 229–257.

Waxman, S. R., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*, 139–156.

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571–581.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Zwicky, A. (1971). In a manner of speaking. *Linguistic Inquiry, 11*(2), 223–233.