Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognition 87 (2003) 179–208
COGNITION
www.elsevier.com/locate/cognit

Infants' understanding of false labeling events: the referential roles of words and the speakers who use them

Melissa A. Koenig[a,*], Catharine H. Echols[b]

[a]*Harvard University, Cambridge, MA, USA*
[b]*The University of Texas, Austin, TX, USA*

Received 22 August 2000; received in revised form 23 September 2002; accepted 2 December 2002

Abstract

The four studies reported here examine whether 16-month-old infants' responses to true and false utterances interact with their knowledge of human agents. In Study 1, infants heard repeated instances either of true or false labeling of common objects; labels came from an active human speaker seated next to the infant. In Study 2, infants experienced the same stimuli and procedure; however, we replaced the human speaker of Study 1 with an audio speaker in the same location. In Study 3, labels came from a hidden audio speaker. In Study 4, a human speaker labeled the objects while facing away from them. In Study 1, infants looked significantly longer to the human agent when she falsely labeled than when she truthfully labeled the objects. Infants did not show a similar pattern of attention for the audio speaker of Study 2, the silent human of Study 3 or the facing-backward speaker of Study 4. In fact, infants who experienced truthful labeling looked significantly longer to the facing-backward labeler of Study 4 than to true labelers of the other three contexts. Additionally, infants were more likely to correct false labels when produced by the human labeler of Study 1 than in any of the other contexts. These findings suggest, first, that infants are developing a critical conception of other human speakers as truthful communicators, and second, that infants understand that human speakers may provide uniquely useful information when a word fails to match its referent. These findings are consistent with the view that infants can recognize differences in knowledge and that such differences can be based on differences in the availability of perceptual experience.

Keywords: Infants; False labeling; Language acquisition; Referential intent

* Corresponding author. Harvard Graduate School of Education, 502 Larsen Hall, Appian Way, Cambridge, MA 02138, USA.
E-mail address: melissa_koenig@gse.harvard.edu (M.A. Koenig).

0010-0277/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0010-0277(03)00002-7

1. Introduction

Words carry information both about the world and about the people who use words. Although there is much evidence suggesting that infants understand that words refer externally to real-world objects, evidence regarding infants' understanding that words reflect the internal life of human speakers is more limited. One place to look for insight into infants' understanding of words as signs of speakers' mental experiences is the domain of verbal mismatches or false labels. Recent findings suggest that during the second year of life infants begin to attribute to others mental states like desire, perception, intention and emotion; however, the understanding that psychological states may differ from and misrepresent observable reality has not been demonstrated until later in development (Meltzoff, 1995; Tomasello & Barton, 1994; Wellman, 1993). The studies reported here, using a combination of both passive (direction of attention, looking time) and productive measures (infants' corrective pointing and labeling), examine possible roots of this understanding by focusing on infants' responses to words that fail to match reality and their treatment of others' false labeling behavior.

Expectations about labels and expectations about speakers who use labels are two issues central to understanding reference in language. These issues can be illustrated, first, by drawing upon important distinctions that are available in semantic theory and, second, by a review of evidence pertaining to these expectations in infancy.

1.1. The internal and external significance of language

Linguists and philosophers of language distinguish between two domains of facts (e.g. see the work of Burge, 1986; Higginbotham, 1989; Jackendoff, 1983; Segal, 1989). One domain is the *external significance of language* (Barwise & Perry, 1983) constituted by facts about the relationships between linguistic expressions and the world we live in and discuss (de Saussure, 1915/1966; Kripke, 1972; Putnam, 1975). Second, there is the *internal significance of language*, facts about the relationship between linguistic expressions and the speakers who use them to form thoughts, communicate ideas, question others and act (Austin, 1962/1975; Grice, 1969; Partee, 1982; Searle, 1969).

1.1.1. The external significance of language

To distinguish true sentences from false sentences, one appreciates, at the very least, that there is a relationship of correspondence, that is, a 'goes together' relationship between linguistic utterances and the world. For example, specific words are referents for specific entities, and a central part of learning these words lies in learning what entities they refer to in one's linguistic community. Thus, an understanding that "cars" refers to [cars], "dogs" to [dogs], and "chairs" to [chairs][1] requires knowing that these words refer to the same real-world entities to which other people refer when using each of these words.

1.1.2. The internal significance of language

A second domain of facts about natural language includes facts about the relations

[1] Brackets ("[]") serve to specify the real-world object denoted by the label.

between language and its speakers. Meaning relates language not only outward to the world of cars, dogs and chairs but also inward to the mental lives of human agents through speaker-oriented notions like saying, believing, asking, doubting, judging, and so on. As explicated below, that one makes inferences about the internal states of others based on what they say is a commonplace observation that adults use implicitly every day. In straightforward cases (e.g. assertion), the meaning of an utterance reflects directly the content of some internal, psychological state. For example, if a person asserts, "That's a chair", it is then likely that her utterance reflects her belief that the object in question is, in fact, a chair and that she wishes to inform some audience that it is a chair.

The relationship between intentional states and speech acts reveals certain similarities as well as differences. First, intentional states 'are about' objects in the same sense that speech acts 'are about' objects (Searle, 1983). Both thoughts and utterances are directed at the world. Second, intentional states contrast with speech acts in terms of the different forms they each take, that is, intentional states are representational states and speech acts are acts or performances. Intentional states are "intrinsically intentional", to be a belief simply is to be "about" something, whereas speech acts are physical entities (e.g. noises from the mouth or marks on paper) that "derive" their intentional properties from the states that motivated them (Searle, 1983, p. 27).

Do infants understand speech acts as related to the intentional states that motivate them? First steps toward an answer to this question come from infants' understanding of words and their use of socio-pragmatic cues in word learning, which we will review below. In light of such considerations, our main experimental questions take the following form. First, are infants' interpretations of the external significance of words influenced by the intentional nature of the source producing the words? Second and relatedly, when word–object links are violated, do infants have different reactions when human agents, as compared to other sources, produce those violations? Third, does this understanding make reference to the internal experiences of human agents who produce word–object violations? Consider an infant who knows what dogs and chairs are called. When confronted with a person who sees a chair and who uses "dog" to refer to a [chair], does that infant demonstrate an understanding that any of these fundamental violations may have occurred: first, a violation of the external correspondence relation between an utterance and a referent; second, a violation of the reputed status of human speakers as truthful labelers; and third, a violation of the internal correspondence relation between an inferred underlying intention and the speaker's utterance?

1.2. Infants' assumptions about words

Research on early word learning and comprehension suggests that infants do expect systematic and veridical relationships between words and objects in the world. The literature on intermodal perception (Kuhl & Meltzoff, 1982; Starkey, Spelke, & Gelman, 1983) suggests that even very young infants prefer to look at visual stimuli that match an auditory stimulus over non-matching visual stimuli. In the realm of language, infants as young as 8–12 months look longer at a familiar object when presented with its corresponding label (e.g. Jackson-Maldonado, Thal, Marchman, Bates, & Gutierrez-Clellen, 1993; Oviatt, 1980). Other research demonstrates that infants will select from a group of objects

that which matches a label (e.g. Benedict, 1979; Nelson, 1973) and that they preferentially orient to an image of an object that previously had been associated with a specific label (Schafer & Plunkett, 1998). At 14 months, infants can learn the association between arbitrary words and moving objects and dishabituate when words are paired with the 'wrong' objects (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Furthermore, if no correct match is available for words they know, 21–27-month-olds fail to exhibit a preference toward any object (Naigles & Gelman, 1995). When 24-month-olds hear a word they know that does not match the object they are looking at, they rapidly shift their gaze to an alternative object and continue to shift gaze in search for the appropriate referent (Swingley & Fernald, 2002).

Further insight into early understanding of the external significance of language comes from work examining young children's expectations about new word–object links. Research on children's initial hypotheses regarding word meaning suggests that young children avoid associating a novel label to an object with a known name, preferring instead to link the novel label to an unnamed object (Golinkoff, Mervis, & Hirsh-Pasek, 1995; Markman, 1989; Markman & Hutchinson, 1984; Markman & Wachtel, 1988; although see Mervis, Golinkoff, & Bertrand, 1994, for arguments against a strong version of this view). One study that derives from this line of work also has implications for children's understanding of word–referent mismatches. In a study by Markman and Wasow (in preparation, cited in Woodward & Markman, 1998), 15- and 18-month-olds were shown an object with a known label – a spoon – in plain view on an otherwise empty table in front of the infant. An experimenter asked, "Can you show me the mido?" Infants searched the room as if to look for a candidate referent for the novel word. Markman and Wasow took such results to provide indirect support for the claim that infants as young as 15 months resist associating second labels with objects whose first labels are already registered in their vocabularies.

Infants' rejection of a second name for a known entity has been attributed to a lexical constraint such as the mutual exclusivity assumption, an expectation that each object will have only one label (Markman, 1989). However, the understanding that leads infants to question "mido" for an object known as "spoon" may not be the same as the understanding that leads infants to reject "dog". In the latter case, infants' knowledge of the meaning of a word likely gives them more information and, thus, further reason to reject "dog" as a possible term for [spoon]; they may take the word (i.e. "dog") to refer uniquely to certain (dog-like) things (Swingley & Fernald, 2002). In the case of false labels, it may be that, in addition to mutual exclusivity, children's expectations regarding the referential role of words has been violated. Moreover, a novel label directed at a spoon might refer to a broader or more specific category of that object (e.g. silverware, soup spoon), but such an interpretation is less plausible with familiar labels (Merriman & Bowman, 1989).

Another contributing factor for why infants reject false labels may be that they have developed an understanding that people use words in certain ways and not others; namely, people use different forms to convey different meanings. Pragmatic explanations of young children's lexical development argue that knowledge of a word's meaning may derive, in part, from the understanding that a person's use of one word, in a given situation, rather than another word reflects an underlying reason or intent (Clark, 1990, 1991). For example, in the Markman and Wasow task, infants may have searched the room for a "mido" on

the assumption that if the experimenter had wanted the spoon, then she would have asked for the spoon, given that "spoon" is the commonly understood term for that particular object. Because the experimenter asked for a "mido", the infant may have reasoned that she wanted something else that contrasts with the meaning of "spoon" (Clark, 1990).

Words that successfully refer to things direct one's attention to objects and offer little opportunity to determine the underlying pragmatic understanding of the labeling events. When presented with familiar (true) labels, infants look longer at the appropriate target object and in doing so reveal some understanding that words refer to things (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Oviatt, 1980). When infants avoid lexical overlap and take a novel word to refer to a nameless object, they further reveal their understanding of the ways words refer to objects (Golinkoff et al., 1995; Markman & Wachtel, 1988; Merriman & Bowman, 1989). Thus, infants' treatment of true labels and novel labels typically shed light on their understanding of how words relate to external objects. However, false labels, or words that fail to refer, could potentially offer the infant the opportunity to express an understanding of words' internal significance. Understanding false labels involves recognizing not only that a word failed to refer to an object but that a person failed to refer correctly.

In sum, it is unclear whether infants' expectations for word–object correspondences are expectations about labels or about human speakers. Most if not all mutual exclusivity research involves the presence of a human experimenter. The word comprehension research mentioned above includes some studies that involve a human experimenter and some studies that do not. Given that infants have succeeded in matching words with objects in the absence of human speakers (Naigles & Gelman, 1995; Schafer & Plunkett, 1998), it may be that infants expect known labels, simply in virtue of their status qua labels-with-meaning, to match up with associated facts in the world. However, because these lines of work have not systematically alternated or compared different labeling sources, we do not know whether the presence of a human speaker may influence infants' interpretations of words they know.

1.3. Infants' assumptions about human speakers

At a fairly young age, infants provide indications that they understand that people use labels with an intent to refer (Baldwin & Moses, 1996; Tomasello & Barton, 1994). Baldwin and colleagues have examined, in a series of studies, what happens in contexts in which adults comment on an object while infants are focused on a different object (Baldwin, 1993a; Baldwin et al., 1996). Results indicate that as early as 16–19 months of age infants determine whether a novel label refers to an object by noticing cues to referential intent, for example, the speaker's line of sight and body posture. This accords with work by Tomasello and Barton (1994), who found that 2-year-olds use behavioral cues indicative of adults' intentionality, such as a joyful facial expression, eye contact with the child, or giving the object to the child, when learning new nouns and verbs. In one study, for example, an experimenter introduced a novel verb (e.g. "Let's dax Big Bird") before producing two novel actions, one accidental and one purposeful. Children in this situation clearly monitored the experimenter's actions, discriminated accidental from intentional actions, and learned the new word for the intended action only. Taken together, these

findings suggest that infants are capable of using multiple cues to infer intent and to determine the object of adult acts of reference.

Research on infants' use of cues to referential intent, like the word-learning research mentioned earlier, has relied on novel labels. This research leaves unanswered questions concerning familiar language use, and thus false language use. It is as yet unclear how infants make sense of utterances that conflict with their experience – what information is attended to when infants recognize this mismatch and whether they make attempts to resolve the conflict.

As we have discussed, adults understand that when someone makes an assertive claim about the world, the statement simultaneously reflects the world as well as that person's belief about that aspect of the world. Although adults may treat straightforward utterances as belief-reports, the nature of 16-month-olds' understanding proves an issue of continued discussion. Beliefs are particularly complex internal states constituted by evaluations, interpretations or critical representations of the world (Bratman, 1987; Searle, 1983) and although infants may understand certain aspects of psychological causality (i.e. goals, intentions, simple desires), the etiology of mental states such as belief may not, as yet, be among them (Astington & Gopnik, 1991; Flavell, 1988; Gopnik, Slaughter, & Meltzoff, 1994; Wellman, 1990).

There are various possible interpretations that infants could have regarding the internal significance of language. First, it seems possible that infants could determine whether a familiar statement refers to a referent without any appreciation of human goals or intentions; they might simply assess whether or not external truth conditions hold for words they know. If the understanding that labels and objects match is dependent solely on associations between words and referents, then in any situation in which words and objects fail to match, we might expect infants to become frustrated and fail to attend regardless of the labeling source. Second, infants may have learned to associate certain entities (i.e. people) with certain human behavior (i.e. accurate labeling in accord with objects). They may have come to expect humans to provide accurate labels but give a non-mentalistic treatment to such behavior, failing to infer that intentions underlie human labeling (Moore, 1999; Perner, 1991). If this is the infants' view, they should treat true human labelers as familiar and false human labelers as novel. However, differences in a speaker's access to information about a potential referent should not influence their interpretation of labels. Third and finally, infants may possess a mentalistic view and understand referential behavior as motivated by the underlying intentional states of people (Meltzoff, 1995; Meltzoff, Gopnik, & Repacholi, 1999; Tomasello & Barton, 1994). If referential failures lead infants to question a speaker's intent, then they should seek to identify an explanation for the mistake by attending to intentional cues of the speaker and indicating an understanding of what the person *meant* to say (e.g. through corrective labeling and pointing). If infants give an intentional interpretation to speech acts, then differences in speakers' knowledge of objects should influence their interpretation of true and false labeling events.

1.4. Motivation for studies

In a series of four studies, we designed situations to determine systematically whether 16-month-old infants' responses to false labels would differ as a function of whether a

human labeler is providing the labels and as a function of the knowledge state of the speaker. We tested 16-month-olds because this is the age at which infants demonstrate a sensitivity to intentional cues in novel labeling contexts (Baldwin, 1993a,b). Half of the infants experienced situations wherein they, for example, looked at a picture of a ball and heard, "Look, that's a ball" (True Condition), and the other half looked at the ball and heard, for example, "Look, that's a duck" (False Condition).

In Study 1, we examined how infants responded to true vs. false human labelers. In this study, the person who was present in the room actively produced either true or false labels. Study 2 tested the effects that a visible, inanimate sound source had on infants' responses to false vs. true labeling events. Labels were produced by an audio speaker situated in the location of the human labeler of Study 1. In Study 3, a control study, we separated the human from the labeling source, such that labels were provided by a concealed audio speaker while the human remained silent. Study 4 tested the effect of a speaker's perceptual access on infants' interpretation of true vs. false labels. In this study, the person actively produced true or false labels while gazing away from the objects. In all four studies, by comparing children's reactions to utterances that corresponded to objects with their reactions to those that did not correspond, we were able to evaluate infants' expectation that veridical relationships hold between words and objects across different contexts. Manipulating across studies the conditions under which a label was provided allowed us to examine infants' expectations regarding humans' labeling behavior.

If infants hold a general expectation that labels refer veridically to the world independently of who or what produces the labels, then across all four studies, a difference in performance between True and False Conditions would be predicted. As in Markman and Wasow (in preparation, cited in Woodward & Markman, 1998), infants could respond to false labeling events by refusing to look at the object and, instead, searching the room. Another possibility is that infants could treat false labeling events as highly novel and thus spend more time looking at them. Given prior research showing that infants tend to look toward or select objects that match a label (e.g. Golinkoff et al., 1987; Naigles & Gelman, 1995), we predicted that infants should spend less time looking to the object during false than during true labeling events. If infants assume that labels qua labels are things that refer veridically, this pattern should appear regardless of the source of the label, and thus should be found in all four studies.

If infants expect that labels match objects when produced by human speakers, and if these expectations are based solely on patterns observed in prior experience with human labelers and do not rely on notions of speaker intent, then infants should respond differently to true and false labels in those studies that include a human labeler (Studies 1 and 4) than in those that do not (Studies 2 and 3). If infants have associated accurate labeling with human speakers, then infants should find false human labelers novel and surprising; in contrast, true human labelers should strike infants as common in their experience and familiar to them. Thus, for Studies 1 and 4, infants should look longer to the human labeler during false than during true labeling. Because Studies 2 and 3 do not involve a human labeler, no such pattern would be expected.

Alternatively, if infants are developing an early understanding of the internal significance of language and expect words to correspond to the perceptual experiences of human speakers, then infants should interpret the human speaker of Study 4 differently than that

of Study 1. If the infant interprets words as referring not only to external objects but also as reflecting the internal states of speakers, then the direction of a speaker's gaze should affect infants' interpretation of true and false labels. If a speaker is gazing directly at an object, infants should infer that the speaker intends to label the object, with the result that a false label elicits attention. Infants also may seek to draw a speaker's attention to her mistake, in which case they would make overt attempts to correct speakers. If the speaker is gazing away from the object, the infant should be less likely to infer an intent to label that object but perhaps express interest in determining the appropriate focus of the speaker's intent. Indeed, in this latter case, when a speaker has no knowledge of the objects presented, true labels should be unexpected and surprising to infants.

2. Study 1

2.1. Method

2.1.1. Participants

The participants were 32 16-month-old infants (age: $M = 16$ months, 7 days, range = 440–470 days) whose names were obtained using birth announcements listed in a local newspaper in Austin, TX. There were equal numbers of male and female infants. All participants were exposed predominantly to English. An additional 12 infants were excluded from the study: six cried or became bored, one was replaced due to technical problems experienced during the study, and five were excluded because they did not comprehend one or more of the target words, based on parental report.

When parents and their infants arrived at the Children's Research Lab, an interviewer went over the fundamentals of the study, obtained informed consent from the parent, administered the vocabulary checklist (which included only those words used in the study) and answered any questions. Parents received a small gift for their infants' participation in the study but were not otherwise compensated.

2.1.2. Stimuli

Across 12 trials, the visual stimuli were a series of photographic color slides of five objects whose labels are familiar to children at this age (Nelson, 1973): chair, duck, cat, ball, shoe. Images were presented one at a time. The words were heard in a standard naming frame, "That's a _____". Two different orders of stimuli were used in each condition: True Order 1, True Order 2, False Order 1 and False Order 2 (see Appendix A).

2.1.3. Set-up

Infants sat on parents' laps. The proximity to the parent helped to ensure comfort with the testing situation. Precautions were taken to ensure that parents did not influence their infants' responses: parents were asked to wear a sleep mask that prevented them from seeing the stimuli and to refrain from talking to the infant during the experiment. The infant and parent were seated approximately 7 feet away from and facing the screen on which the slides were projected. A person, who sat in a chair 2–3 feet to the left of the infant, actively labeled the objects while gazing at the stimuli (see Fig. 1). In an effort to

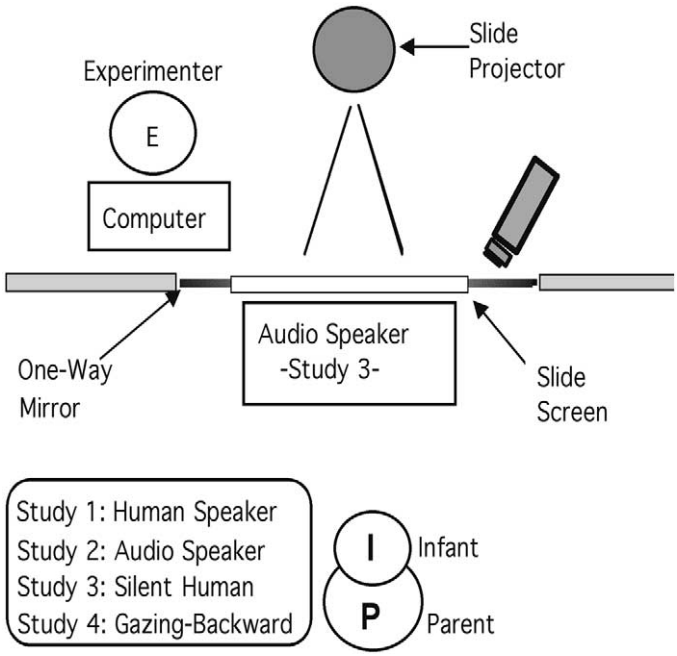

Fig. 1. Experimental set-up.

ensure that labeling episodes were similar for all participants, the human speaker did not engage in one-on-one interaction with the infant. Instead, the human speaker looked attentively ahead at the screen upon which the objects were projected and labeled the objects. The human speaker uttered labels using a standard naming frame, for example, "That's a dog", "That's a cup".

2.1.4. Design and procedure

Half of the participants heard familiar words that correctly labeled the familiar objects and half heard familiar words that incorrectly labeled the familiar objects. For example, infants in the True Condition saw a picture of a ball on the screen and heard, "That's a ball". Infants in the False Condition, for example, saw a picture of a ball and heard, "That's a shoe". To maintain attention, the human speaker repeated the utterance every 4 s while the image was displayed. The beginning and ending of each trial was controlled by a concealed experimenter who advanced the slides. Each image remained until the infant looked away from the scene (either the screen or the person) for a period of 2.5 s or until 30 s had elapsed. The infants' eye movements were recorded using a camera located behind a one-way reflective window.

2.1.5. Coding

The data that were submitted to analysis were coded from videotapes recorded during the experiment. All videotapes were coded independently by the primary investigator and

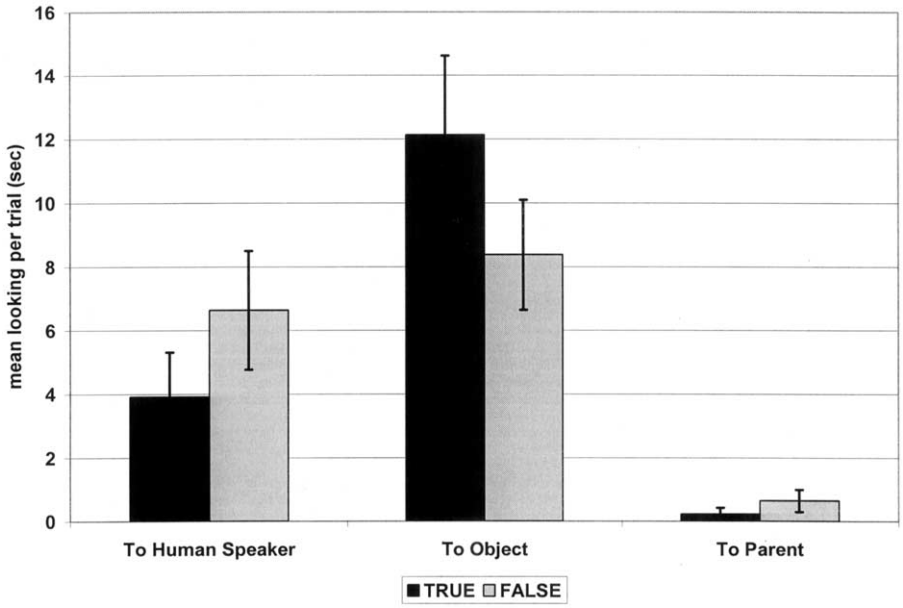

Fig. 2. Infants' looking behavior during true and false trials for Study 1.

two undergraduate research assistants. Separate patterns of looking were determined for infants' looking to the object, to the observer, and to the parent. Behaviors that were coded included (1) total looking time to object during each trial, (2) total looking time to observer for each trial, (3) total looking time to parent for each trial, and (4) infants' pointing and labeling behavior (i.e. number of times infants vocalized or pointed to the object and where they were looking when they did so). Percent agreement between the two coders over a total of 516 cases/trials ranged between 91% and 97% on the dependent looking measures. Overall reliability using Cochran's test (Cochran, 1950) was $Q = 0.94$, $M = 9.01$–9.27, $N = 516$, $P < 0.05$.

2.2. Results

Given that the looking behaviors toward the object, human and parent were recorded as dependant measures for each subject, mean looking times were calculated across 12 trials and the data from Study 1 were entered into a 2 (slide orders: 1, 2) × 2 (condition: true, false) MANOVA with order and condition as between-subject factors and with mean looking to the object, human speaker and parent as dependent measures.[2]

The MANOVA revealed a main effect for condition ($F(3, 26) = 6.85$, $P < 0.001$). There was no effect for order nor was there a condition × order interaction. Mean looking

[2] Looking to object precluded looking to the human and the parent at any given point in time but because maximum looking to object was 30 s and infants rarely looked for the entire 30 s, they had ample time to reference each of these different fixation points. The issue of possible trade-offs between looking to the object and labeling source is addressed more fully in Section 7.

times for each of the dependent measures are shown in Fig. 2. For the dependent measure of looking to object, the main effect of condition was significant ($F(1, 28) = 6.10$, $P < 0.05$), with infants looking longer to the object during true than during false trials. In contrast, infants looked significantly longer to the human speaker during false than during true labeling trials ($F(1, 28) = 5.13$, $P < 0.05$), and to the parent during false than true trials ($F(1, 28) = 4.52$, $P < 0.05$).

2.3. Discussion

In Study 1, infants presented with false labeling looked significantly longer to the human speaker and to the parent than did infants who were presented with true utterances. This is a notable finding for several reasons. Labeling activity, accompanied by joint attention, might have engaged infants' attention and increased attentiveness quite generally to the human, who was the source of the sounds, with the result that attention to the human would have been consistent across both true and false trials. Alternatively, because the labels infants encounter from people are more typically true than false, the familiarity of true labeling might have led infants to pay more attention to the human during that condition. The actual findings are consistent with the prediction that infants expect humans to provide true labels, thus are surprised by false labels, and perhaps motivated to resolve the conflict by accessing cues provided by human speakers.

This study also showed that the amount of attention infants gave to the object differed for true as compared to false utterances. Infants looked longer at the object when it was correctly labeled than when it was incorrectly labeled. This finding corresponds with the research on lexical development suggesting that infants prefer that labels match the object with which they have been associated. A remaining question concerns whether or not the observed patterns of looking behavior were affected by the fact that the utterances were communicated by a human agent. To examine whether infants' responses to true and false utterances interact with their understanding of the source of the labels, in Study 2 the true and false labels were articulated by an inanimate speaker.

3. Study 2

3.1. Method

3.1.1. Participants

The participants were 32 16-month-old infants (age: $M = 16$ months, 7 days, range = 440–470 days). Again, there were equal numbers of male and female infants. An additional 13 infants were excluded from the study: four cried or became bored, one was replaced due to technical problems experienced during the study, and eight infants were excluded because parental report indicated that they did not comprehend one or more of the target words. Participants for Study 2 were recruited and excluded according to the same considerations as those of the first study.

3.1.2. Stimuli and procedure

Study 2 incorporated the same stimuli and procedure as Study 1. The only difference

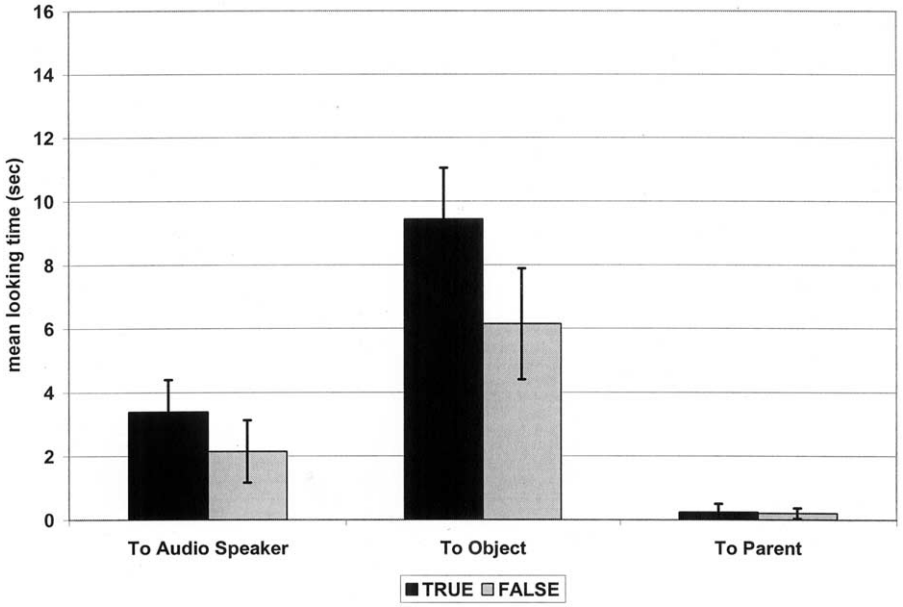

Fig. 3. Infants' looking behavior during true and false trials for Study 2.

that distinguished the protocol of Study 2 from Study 1 was this: in Study 2, a Sony audio speaker sat in the chair to the left of the infant and emitted the labels for the objects (see Fig. 1). Labels and objects were presented using the orders of Study 1 and the same naming frame (i.e. "That's a _____").

3.1.3. Coding

The coding procedures were the same as those used in Study 1.

3.2. Results

As for Study 1, the data from Study 2 were entered into a 2 (slide orders: 1, 2) × 2 (condition: true, false) MANOVA with order and condition as between-subject factors and with mean looking to the object, audio speaker and parent as dependent measures. The MANOVA indicated a main effect for condition ($F(3, 28) = 3.85, P < 0.01$). Fig. 3 shows mean looking times for each of the dependent measures. As in Study 1, infants looked longer to the object during true than during false labeling ($F(1, 30) = 7.73, P < 0.01$). In contrast to Study 1, infants did not look longer to the audio speaker during false than during true trials. In fact, we found the following unexpected non-significant tendency: infants tended to look longer to the audio speaker during true than during false labeling ($F(1, 30) = 3.14, P = 0.086$). No effects were found for condition on infants' looking to parent.

3.3. Discussion

Infants' expectations that familiar or learned labels match the objects with which they have been associated seem also to coincide with a developing expectation in infants that humans generally are reliable and truthful labelers. In Study 1, a human agent labeled the objects; in this context, infants referenced the human speaker more during false than during true labeling. Study 2 provided a context in which the source of object labels was not human; infants did not reference the source of false labels in this context. When linguistic utterances conflicted with infants' experience, they attended to a human speaker responsible for the statement but did not attend to non-human sources. These results suggest that infants have different expectations about human and inanimate labelers.

Consistent with our predictions, with results of Study 1, and with work previously done in lexical development, infants in the False Condition of Study 2 were significantly less interested in the labeled objects than were infants in the True Condition. This general pattern of results toward the object did not seem to differ as a function of whether an audio speaker or a human speaker was responsible for the labeling. These results confirm that 16-month-old infants expect words to map consistently and appropriately to a referent and that their preference for matching word–object pairs does not depend upon the animacy of a labeling source.

Although infants' attention to the human speaker during false labeling is consistent with the possibility that they have certain expectations about human behavior, it is possible that infants looked longer to humans of Study 1 for another reason. It could be that infants' attention to the false human labeler is an artifact of their aversion to attend to falsely labeled objects. It could be argued that because humans are more interesting than audio speakers, infants' attention away from falsely labeled objects was directed at the next most interesting thing present – a human agent. To rule out the possibility that infants' differential attention to the labeling source in Studies 1 and 2 is a result of infants' interest in humans over audio speakers, in Study 3, a human being sat in the same position as that of Study 2's human speaker. The utterances were presented by a concealed audio speaker positioned near the visual presentation of objects, similar to what a child might experience in the context of a television. In this study, if infants show little interest in the human during the False Condition, we could get closer to inferring that infants' attention to the false labeler in Study 1 is due uniquely to the false labeling behavior of the human.

4. Study 3

4.1. Method

4.1.1. Participants

The subjects were 32 infants (16 of each gender) who were approximately 16 months old (age: $M = 16$ months, 12 days, range = 440–470 days). An additional 15 infants were excluded from the study: three cried or became bored, two were replaced due to technical problems experienced during the study, and ten infants were replaced for failure to know

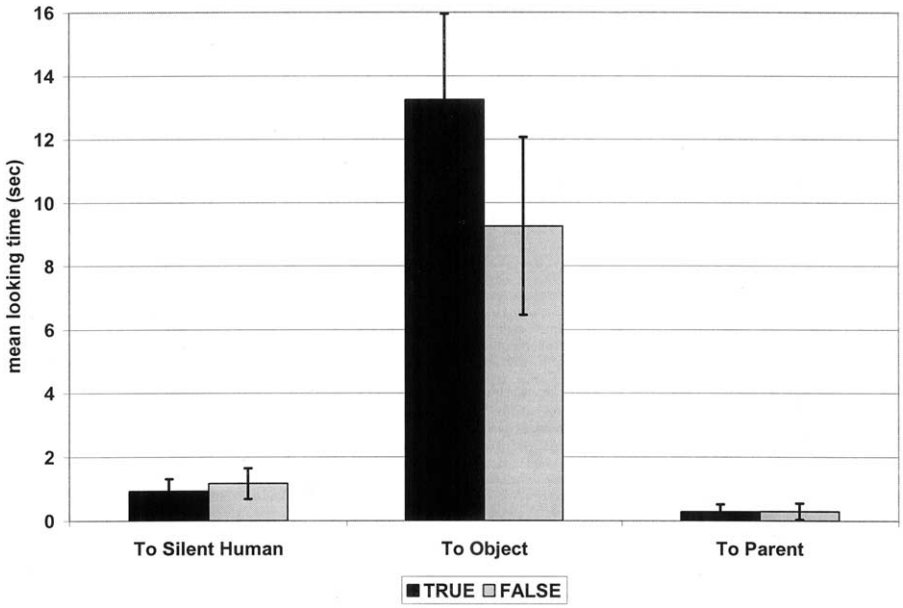

Fig. 4. Infants' looking behavior during true and false trials for Study 3.

one or more of the target words. Participants for Study 3 were recruited and excluded according to the same considerations as those of the first two studies.

4.1.2. Stimuli and procedure

Study 3 incorporated the same stimuli and procedure as Studies 1 and 2. The only difference that distinguishes the protocol of Study 3 from the first two is this: the audio speaker (which sat near the infant in Study 2) was hidden under a table that was directly in front of and below the screen (see Fig. 1). An observer, who sat silently 2–3 feet to the left of the infant, did not engage in any interaction with the infant as she watched the same stimuli as the infant. The audio speaker emitted the labels in the same naming frame used in the first two studies (i.e. "That's a _____").

4.1.3. Coding

The coding procedures were the same as those used in Studies 1 and 2.

4.2. Results

The data from Study 3 were entered into a 2 (slide orders: 1, 2) × 2 (condition: true, false) MANOVA with order and condition as between-subject factors and with mean looking to the object, silent human and parent as dependent measures. There was a marginally significant effect of condition for infants' looking to the objects ($F(1, 30) = 4.14$, $P = 0.051$), with infants tending to look longer to the object for true than for false utterances (see Fig. 4). There was no effect of condition on infants' looking

to either the silent observer or the parent. In fact, looking times toward the silent observer and parent generally were quite low, with overall means of 0.93 and 0.28 s, respectively, for true trials and 1.16 and 0.3 s for false trials.

4.3. Discussion

In Study 3, although infants tended to look away from the object longer during false than during true labeling, they did not redirect their attention toward the silent human agents – observer or parent – during false trials. In the context of multiple mismatches (i.e. false utterances), infants tended not to refer to the human agents present in the room. Thus, infants' attention to the false labeling source of Study 1 cannot be fully explained as an artifact of their aversion to falsely labeled objects. Instead, infants' attention to the false labeling source of Study 1 derived from the fact that a person was responsible for false utterances.

With regard to the issue of infants' understanding of the referential role of words, we find two observations worthy of mention. First, across three different labeling contexts, infants who saw objects that were correctly labeled generally looked longer at objects than infants who saw objects labeled incorrectly. This finding corresponds with the research on lexical development which shows that infants prefer that labels match the objects with which they have been associated and suggests further that infants' expectation that labels refer veridically holds independently of who or what produces the labels. A second observation pertains to infants' interpretations of the mismatch. Speculatively, if infants took object–label mismatches as evidence that they had been mistaken about a word's referent, one might expect them to show heightened looking to the object in an effort to rectify the mistake and learn the correct label. Infants' decreased looking to the object in the False Condition hints that they do not interpret the false labels as evidence of an error on their part. Consistent with this interpretation, infants never repeated or rehearsed false labels, though they repeated true labels (see discussion of points and labels in Section 6.2).

In sum, infants' attention to the labeling sources in these three contexts suggests that, when linguistic utterances conflict with infants' experience, infants' interpretations are sensitive to the nature of the information source responsible for the behavior. However, as discussed earlier, the nature of infants' understanding is open to at least two possibilities, a non-intentional and an intentional understanding. It may simply be that infants have learned to associate a certain type of behavior, linguistic reference, with certain types of entities observed to have referential capacity, people. People tend to talk about things that are relevant and consistent with infants' experience. When there is a mismatch between a person's label and a referent, infants' expectations are violated and result in their surprise. Alternatively, infants may respond with surprise toward false human labelers because of attributions they make regarding others' intentional states. Infants may be surprised to observe false labels from someone who has perceptual access to objects and reason to label truthfully. Study 4 is designed to test this possibility by presenting infants with a true or false human labeler who faced backward and gazed away from objects. If infants' attention to false labelers is simply due to a violated norm that links humans with correct labeling, we would expect infants to look longer to the false than to the true labelers. However, if the results of Study 1 were due to an early intentional understanding,

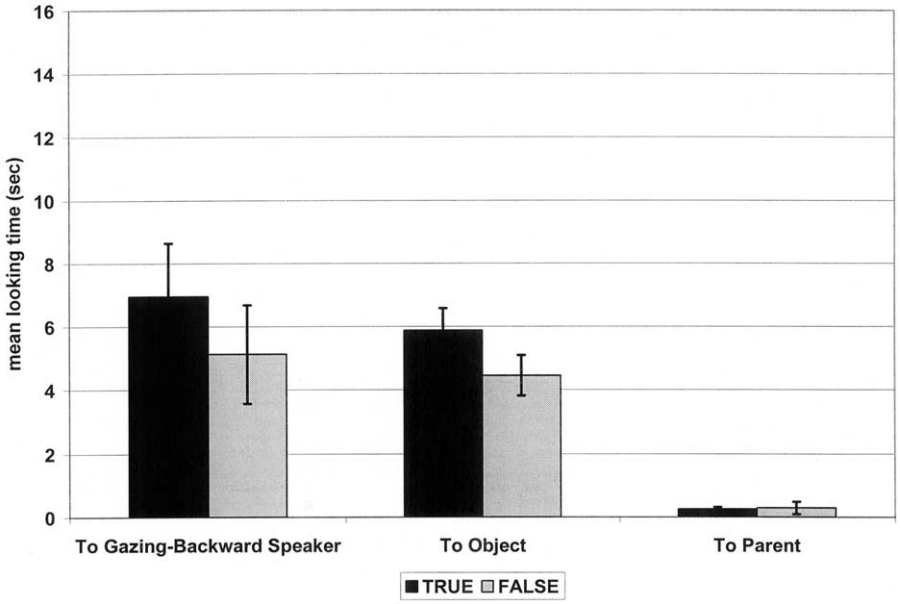

Fig. 5. Infants' looking behavior during true and false trials for Study 4.

we would expect to see a difference between infants' behavior toward facing-backward (Study 4) and gazing-toward human labelers (Study 1).

5. Study 4

5.1. Method

5.1.1. Participants

The participants were 32 16-month-old infants (age: $M = 16$ months, 4 days, range = 440–470 days). Again, there were equal numbers of male and female infants. An additional 16 infants were excluded from the study: six cried or became bored, three were replaced due to technical problems experienced during the study, and seven infants were replaced whose vocabulary checklists indicated that they did not know one or more of the target words. Participants for Study 4 were recruited and excluded according to the same considerations as those of the first three studies.

5.1.2. Stimuli and procedure

Study 4 incorporated the same stimuli and procedure as Studies 1, 2 and 3. The only difference that distinguishes the protocol of Study 4 is this: in Study 4, a human speaker sat in the chair to the left of the infant with her back facing the objects and produced the labels for the objects (see Fig. 1). A small audio speaker placed inside the person's ear was

connected to the experimental computer, which relayed the appropriate label to the human speaker. The labels were presented in two different orders in the same naming frame used in the first three studies (i.e. "That's a _____").

5.1.3. Coding

The data collection and coding procedures were the same as those used in Studies 1, 2 and 3.

5.2. Results

The data from Study 4 were entered into a 2 (slide orders: 1, 2) × 2 (condition: true, false) MANOVA with order and condition as between-subject factors and with mean looking to the object, human speaker and parent as dependent measures. Means for each of the dependent measures are shown in Fig. 5. In contrast to Studies 1, 2 and 3, in which infants looked longer to objects during true than false labeling, the effect of condition on the dependent measure of looking to object was non-significant in Study 4 ($F(1, 30) = 2.18$, NS). Regarding infants' looking to the facing-backward human speaker, infants did not look longer to the human during false than during true trials as they did in Study 1. In contrast, infants tended to look longer to the gazing-backward labeler during true than during false labeling ($F(1, 30) = 3.67$, $P = 0.065$). No effects were found for condition on infants' looking to parent.

5.3. Discussion

Study 4 was designed to test whether infants' attention to the false labeler of Study 1 was due to the violation of an expected association between humans and truthful labeling. If infants expect humans to label objects truthfully because they have learned to associate a certain object category (humans) with a certain event category (accurate labeling), then in the False Conditions of Studies 1 and 4, in which humans fail to label objects successfully, infants should express surprise and orient to the false human labeler. However, if infants' interpretation of labeling makes reference to the intentional states of speakers, then infants' interpretation of labeling events will be sensitive to the different perceptual experiences of the labelers in Studies 1 and 4. Consistent with the view that infants have an intentional understanding of behavior, infants did not reference the false human labeler more than the true human labeler when they faced away from objects. In contrast, under this gazing-away condition, infants tended to look longer at the truthful labeler. An independent samples *t*-test indicated that infants looked longer at the gazing-away true labeler ($M = 6.95$) than the gazing-forward true labeler ($M = 3.92$) ($t(30) = 2.91$, $P < 0.01$). Infants may have been surprised that a person with no perceptual experience nor knowledge of an object would be able to label it correctly.

Infants of Study 4, in contrast to three prior studies, did not look longer to the object during true labeling. In fact, post-hoc tests revealed that infants' attention to the object generally was lower in Study 4 than Study 1 (Tukey HSD ($n = 64$), SE = 1.02, $P < 0.001$), non-significantly lower than Study 2 (Tukey HSD ($n = 64$), SE = 1.02, $P = 0.055$), and lower than Study 3 (Tukey HSD ($n = 64$), SE = 1.02, $P < 0.001$). Mean looking times per trial to the objects in the True vs. (False) conditions for Studies

1 through 4 were 12.13 (8.37), 9.43 (6.14), 13.23 (9.26), and 5.87 (4.45), respectively. The absence of a true–false difference in looking to objects in Study 4 suggests that infants did not necessarily accept the labeling behavior of a person who was gazing backward to be relevant to the objects they saw in the front of the room. Infants' preference for correctly labeled objects over falsely labeled objects does not apply when referential cues (gaze and labeling) are in conflict, i.e. gaze directed one way and labeling in another.

6. Between-study comparisons

Given that our main question of interest concerns infants' understanding of different types of labelers, the analyses that follow include between-study comparisons of infants' looking behavior. We also include analyses of infants' pointing and labeling behavior. Analyses of looking behavior alone, without more productive measures, can render it very difficult to tease apart empirically an understanding of intentions vs. action-based aims in infancy (Baron-Cohen, 1994; Meltzoff, 1995; Woodward, 1998). For purposes of simplification, the various entities situated to the left of the infant (i.e. the human speaker of Study 1, the audio speaker of Study 2, the silent observer of Study 3 and the gazing-backward human labeler of Study 4) will be described henceforth as the "leftward entity".

6.1. Looking behavior

To further clarify and compare the results of the experiments, the data were entered into a 4 (Study: 1, 2, 3, and 4) × 2 (condition: true, false) MANOVA with study and condition as between-subject factors and with mean looking to the object, leftward entity and parent as dependent measures.

First, we found a significant Study × Condition interaction for looking to the leftward entity ($F(3, 120) = 5.78$, $P < 0.001$). To better understand this interaction, three separate analyses were performed between Study 1 and each of the other three studies. For each two-way comparison, a 2 (Study) and 2 (Condition) MANOVA revealed a significant Study × Condition interaction for looking to the leftward entity: Study 1 vs. Study 2, $F(1, 60) = 8.52$, $P < 0.01$; Study 1 vs. 3, $F(1, 60) = 4.29$, $P < 0.05$; Study 1 vs. Study 4, $F(1, 60) = 9.14$, $P < 0.005$. The difference in looking toward the leftward entity during false as compared to true labeling was greatest for the human speaker who gazed toward the objects while labeling them (i.e. Study 1).

A main effect for Study was reflected in infants' looking to the leftward entity: the active, talking human speakers of Study 1 and Study 4 attracted more attention from infants than did the silent human of Study 3 (Tukey HSD ($n = 64$), SE = 0.596, $P < 0.001$; Tukey HSD ($n = 64$), SE = 0.596, $P < 0.001$) or the audio speaker of Study 2 (Tukey HSD ($n = 64$), SE = 0.596, $P < 0.001$). That is, regardless of the direction of the speaker's gaze, infants found the behavior of human labelers (Study 1: $M = 5.27$, SD = 3.52; Study 4: $M = 6.04$, SD = 2.81) more interesting than that of the leftward entities of Study 2 ($M = 2.77$, SD = 2.06) and Study 3 ($M = 1.05$, SD = 0.85).

A third and final result was a main effect for Condition in infants' looking to the object: infants looked significantly longer to the object during true than during false labeling

($F(1, 120) = 18.49$, $P < 0.0001$). No main effect for Condition was observed for infants' attention to the leftward entity or to parents.

6.2. Points and labels

Out of 128 total participants, approximately half of the participants did not engage in any pointing or labeling. However, a total of 74 infants attempted to engage in some kind of ostensive activity that appeared relevant to the labeling episodes.

A total of 45 (23 in the True Condition, 22 in the False Condition, respectively) infants pointed to objects: 15 (7 in True, 8 in False) in Study 1; 10 (4, 6) in Study 2; 9 (4, 5) in Study 3; and 11 (8, 3) in Study 4. Sixty-four infants produced labels for the objects: 24 (9 in True, 15 in False) in Study 1; 14 (7, 7) in Study 2; 11 (6, 5) in Study 3; and 15 (7, 8) in Study 4. To better understand the behavior of pointers and labelers in these studies, we coded various productive behaviors including where infants were looking when they pointed and labeled (i.e. object, leftward entity or parent). Four mutually exclusive dependent measures of infants' productive behaviors were coded. They included: (1) infants' non-labeling activity (i.e. points and/or vocalizations); (2) infants' labels; (3) infants' attempts to remove parents' blindfolds; and (4) infants who cried (but recovered). Data were entered into a 2 (condition: true, false) × 4 (Study: 1, 2, 3, 4) MANOVA with condition and study as between-subjects factors and with infants' non-labeling activity, infants' labels, infants' attempts to remove blindfolds and infants who cried as dependent measures. Results indicated main effects for Study for infants' pointing behavior ($F(3, 120) = 2.71$, $P < 0.05$) and for infants' labels ($F(3, 120) = 3.05$, $P < 0.05$). Infants produced more communicative gestures and utterances in Study 1 than in Studies 2, 3 and 4.

Regarding 29 infants' labeling attempts in the True Conditions, across the 168 times it occurred, verbal communication consisted of repetitions of object names consistent both with what they heard and with what they saw. In the False Condition, across all studies, 33 infants produced labels and, interestingly, did not repeat the false labels they heard. Instead, when infants produced labels in the False Condition, a total of 135 times, they produced the correct label for the object they saw. The only exception to this trend were two infants who repeated false labels in Study 4, in which the labeler was facing backward, permitting the interpretation that the labeler's intended referent was an alternative object. More infants (15 out of 16 infants) correctively labeled in Study 1, when a human gazing at the objects was responsible for the mistaken label, than in the False Conditions of either Studies 2 (7/16), 3 (6/16 infants) or 4 (8/16 infants) ($\chi^2(1) = 12.25$, $P < 0.001$). In contrast, no effect for Study was found for infants' repetitions of true labels; infants were as likely to engage in labeling in the True Conditions of all four studies. Infants in the False Conditions of all four studies were significantly more likely to try to remove their parents' sleep mask than were infants of the True Condition ($F(1, 120) = 4.15$, $P < 0.05$). For infants in Study 4 only, a non-significant trend toward an effect for Condition was found in the number of infants who pointed; there was a tendency for more infants in the True than the False Condition to point to objects ($t(30) = 1.91$, $P = 0.06$). A non-significant trend toward a Study × Condition interaction was observed for infants' crying beha-

vior ($F(3, 120) = 2.49$, $P = 0.06$). Infants in the False Condition of Study 1 showed a tendency to become distressed more frequently than infants in any other condition.

7. General discussion

The present studies examined the effect of different labeling sources on infants' interpretation of false vs. true labeling events. Specifically, our experimental paradigm allowed us to investigate infants' attention toward labelers as well as to the object during true and false labeling events. Our appeal to infants' spontaneous productions, particularly their corrections during false trials, allowed us to examine the conditions under which infants would imitate or correct what they heard. Our major novel findings were (1) false labels elicited heightened attention from infants when the labeling source was human but not when it was non-human, (2) infants' interpretations of true and false labeling events were influenced by differences in the speaker's informational access and (3) when false labels derived from human agents gazing at objects, nearly all infants made explicit attempts to interrupt and repair assertive statements. Such findings indicate that truthful labeling is expected and sought from human labelers particularly when speakers have visual access to objects. In addition, infants' tendency to look longer at objects during true than during false labeling across the contexts of the first three studies is an extension of prior evidence found in research on early lexical development (e.g. Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1996; Jackson-Maldonado et al., 1993). In that this pattern was found across systematic variations in pragmatic support, our results are an extension of prior findings.

As discussed, analyses of looking behavior have established that when confronted with a falsehood, infants look at a present, human speaker more than they look at silent humans or inanimate speakers placed in the same location. The implications of infants' responses to false labels could reflect different understandings ranging in their degree of sophistication. Possible interpretations of results ranging from the least to most sophisticated are as follows: first, infants may prefer to look at correctly labeled objects and when those are unavailable, they may prefer talking humans; second, infants may have learned through experience that humans tend to use accurate labels, thus expecting true labeling from humans but not non-humans and being surprised when this expectation is violated; third, infants may have learned that humans provide useful cues to referents but not view these cues as reflections of speakers' intent; fourth, infants may interpret labels as reflecting speakers' intent.

Regarding the first possibility, although infants did give ample attention to correctly labeled objects, our results cannot be explained by a hierarchy of infant preferences with a preference for talking humans ranking second to their preference for correctly labeled objects. If this were the case, we would expect infants' looking time to objects to be lowest in the False Condition of Study 1; however, this was not the case. In fact, for infants in Study 1's False Condition, looking to falsely labeled objects ($M = 8.36$ s) did not differ from looking to the active human speaker ($M = 6.63$ s) ($t(15) = 2.04$, NS). Furthermore, for the human labeling context of Study 1 and the non-human labeling context of Study 2, post-hoc tests indicated no differences in infants' attention to falsely labeled objects; the only significant variation in looking during false trials occurred for attention to the left-

ward entity. Infants looked longer to both the object and the speaker in the naturalistic naming context of Study 1 than in the other three studies ($F(3, 120) = 6.50$, $P < 0.001$), suggesting that results are not just a matter of a trade-off between looks to object vs. the leftward entity. Moreover, it is important to note that amounts of looking to the object during True and False Conditions can be coupled with less *or* more looking to the leftward entity as evidenced by the low, flat looking to the silent observer during both conditions of Study 3 vs. the generally higher level of looking to the audio speaker during true labeling in Study 2.

According to the second interpretation, infants have developed associations between a type of action, true labeling, and the types of objects observed to have referential capacity, people. Thus, people usually tend to talk about things that are available to infants whereas inanimate things like stereo speakers, although capable of emitting words, do not talk about things that are the focus of infants' attention. Such behavioral sequences may become familiar events to infants and if violated, would become unfamiliar or novel, thus eliciting increased attention. This interpretation finds partial support in the finding that, for infants of the False Condition, attention to the human labeler of both Studies 1 and 4 was higher than infants' attention to the leftward entities of Studies 2 and 3. Infants may have learned to associate accurate labeling with talking humans but not with non-human information sources; thus, only when people actively produced false labels were infants' expectations violated. However, note that if infants' surprise toward false labelers was primarily a result of simple associations between humans and truthfulness, we would expect infants to look longer to false labelers than true labelers of both Studies 1 and 4. In contrast, infants' heightened looking to false human labelers relative to true human labelers was obtained only for Study 1, not Study 4. Nonetheless, consistent with an associationist account, it is also interesting to note infants' tendency, albeit non-significant, in Study 2 to look longer to the audio speaker during true than during false utterances. It could be that reliable labeling is behavior they expect from animate beings, with eyes, ears and signs of an observant mental life, and that veridical and reliable labeling from entities that lack these relevant attributes violates their expectations of inanimate entities like stereo speakers. Perhaps infants would respond differently to the contingent, labeling behavior of stereo speakers if facial characteristics were placed on the box. Further research that systematically manipulated the animacy cues of the labeling entity would constitute steps towards answers to such questions.

The third interpretation, mentioned above, is that infants are looking to the human to seek clarifying information. This interpretation, which does not necessarily involve an appeal to intentional states, is consistent with the enhanced looking to the human speakers of Studies 1 and 4 across both True and False Conditions ($F(3, 120) = 29.63$, $P < 0.0001$). In accord with arguments endorsed by Baldwin and Moses (1996), infants' efforts to communicate might be attempts to elicit useful information from other human beings. Consistent with this possibility, we found that infants were more likely to engage actively, by vocalizing or pointing, in the context of the human speaker of Study 1 than the contexts of the other three studies ($F(3, 120) = 11.04$, $P < 0.0001$). These vocalizations consisted of infant articulations ranging from mumbles to queries like "dat?" and "mmm?" and various other attempts to engage with the speaker. Moreover, during such activity, infants were more likely to direct their focus to the labeling source in the human labeling

contexts of Studies 1 and 4 than the contexts of Studies 2 and 3 ($F(3, 120) = 3.44$, $P < 0.02$). Interestingly, in what appeared to be attempts to elicit help from parents in bizarre situations, infants in the False Conditions of Studies 1, 3 and 4 were significantly more likely to try to remove their parents' sleep mask than were infants of the True Condition. In contrast, infants presented with the audio speaker never removed their parents' sleep mask. In sum, infants' interest in humans in these contexts is in line with research suggesting that as infants develop language skills and consume information from the world, they do so with a basic awareness that people provide information that is relevant to the communicative context (Baldwin & Moses, 1996; Bates, Camaioni, & Volterra, 1975). When the informativeness of human labelers is blocked or confused, infants appear able to employ different means to seek cues that might clarify or potentially explain word–object mismatches.

The fourth interpretation mentioned above is that infants take labels to reflect the intentional states of human speakers. One source of evidence for this interpretation derives from infants' corrective labeling behavior. Two-year-olds have been observed to repair and tailor their own unclear utterances and to reject experimenters' false statements (Pea, 1980; Shatz & O'Reilly, 1990; Tomasello, Farrar, & Dines, 1984); however, to our knowledge, very few experiments have systematically investigated infants' ability to repair the mistaken assertions of others. It has been argued that the repair of assertions is particularly valuable in indicating a more sophisticated understanding of others given that the repair of other speech acts such as commands and requests could be explained by infants' persistent abilities to satisfy their immediate needs (Shatz & O'Reilly, 1990).

Several findings from these analyses support this final possibility. First, although the 16-month-olds of our study were not producing words with perfect pronunciation, an examination of the content of infants' labeling attempts proves illuminating when one compares infants of the True and False Conditions. Recall that infants' labeling attempts in the True Conditions consisted of repetitions of object names consistent both with what they heard and with what they saw but that, in contrast, the infants of the False Conditions who produced labels generally did not repeat the false labels they heard. Instead, when infants attempted to produce labels in the False Conditions, they tried to produce the correct label for the object they saw. Most importantly, not only did infants of the False Condition consistently attempt the "correct answer", they did so more often in Study 1: more infants (15/16) correctively labeled when a human with access to objects was responsible for the mistake than in the False Conditions of either Studies 2 (7/16), 3 (6/16) or 4 (8/16). In contrast, infants were equally likely to engage in labeling in the True Conditions across all four studies. It seems likely that infants' corrections, when addressed to human speakers with perceptual access to objects, aimed to draw attention to a person's mistake in order to achieve a shared understanding of reality by conversational means. The smaller number of such corrective attempts during false labeling in Study 4 suggests that infants did not attribute an intent to label objects to someone who lacked perceptual access and who thus had no knowledge of objects.

Further support for this interpretation is that infants were more likely to both point and vocalize while alternating their gaze between the object and the leftward entity when false labels came from a human speaker than when they came from a stereo speaker or when labels were true ($F(2, 48) = 4.11$, $P < 0.05$). Infants' shifting object-to-human gaze

pattern while talking and pointing in the face of word–object discrepancies conforms with prior findings suggesting that infants will shift their gaze from the object to the addressee to supplement the force of their own messages (O'Neill, 1996). Specific behavioral responses from the 16 infants who heard false labels in Study 1 included one infant shaking her head, three infants waving their hands, and five infants cleverly pointing to their own shoe for the human speaker who said, "That's a shoe" but was looking at a [cat], for example. This last behavior may signal an understanding in infants of what human speakers apparently 'meant' to say but, for some reason, did not say. It remains possible that infants' corrective attention to adults' mistakes may reflect attempts to ensure that humans achieve a goal of veridical labeling because this is what humans tend to do when labeling objects yet not reflect an inference of internal states. Whether or not it is motivated by infants' inference of a mistaken intentional state, such an understanding would constitute an attribution to others of a fairly abstract and useful goal: truthful reference.

However, the mentalistic interpretation finds further support in the finding that infants in the True Condition looked longer to the true labeler of Study 4 who lacked access to objects than to true labelers of the other three studies ($F(3, 60) = 17.79$, $P < 0.0001$). Furthermore, a non-significant tendency for more infants to point during true than during false labeling in Study 4 may indicate that infants were particularly interested in directing the true labeler's attention to the objects at which that individual did not, but could have, looked. Such results suggest that a human labeler who achieved truthful reference without access to referents surprised infants. This suggests that infants' interpretation of familiar labeling events depends not only on whether labels match referents but whether the accuracy of a person's labeling behavior can be explained appropriately by that person's access to relevant information.

As discussed earlier, there are certain properties that words and beliefs have in common (Searle, 1983). The first specifies that both words and beliefs are formed to appropriately 'fit' or describe states of the world. Previous research as well as the looking behavior of infants in the current studies suggest that infants construe talking or labeling in terms of actions that are 'directed at' referents. In fact, infants spent more time looking at true referents than at true labeling sources. Another property shared by words and beliefs is the notion of "satisfaction": both false labels and false beliefs constitute failures or mistakes that require change. Note that infants could have expressed surprise by giving heightened attention to the objects in the False Condition of Studies 1 and 4. Instead, infants expressed surprise in the direction of the human agent, suggesting that they attributed the failure to the statement and the human source of the statement, not the world.

What cues do infants use to assess referential acts? Body posture and line of regard, as well as the presence of eyes, mouth and ears, were among the behavioral cues that distinguished human speakers from audio speakers. Behavioral signals of perception such as gaze direction seem an important source of information used by infants and adults to guide an understanding of what other people do. Work by O'Neill (1996) suggests that 2-year-olds appreciate that their mother cannot act, without the aid of attenuated verbal instructions, to retrieve an object she has not seen. Research by Baldwin (1991) demonstrates that 16–19-month-olds learn a novel object's name if the speaker looks at the object while speaking but not if she gazes elsewhere. Infants evidently take the direction of the speaker's gaze as relevant information regarding what the person is talking about. Work

by Spelke, Phillips, and Woodward (1995) provides evidence that by 12 months, infants make inferences regarding a person's behavior from information they have about line of regard and emotional expression. In sum, infants' inferences about a person's actions seem guided by behavioral indications of what the person sees and wants.

Other evidence that infants understand the communicative aspect of human interaction makes appeal to infants' emerging behaviors: infants' gestures become increasingly ritualized with age (Bates et al., 1975), they persevere in the face of failed message transmissions (Bretherton, McNew, & Beeghly-Smith, 1981), they become increasingly astute in repairing their original messages (Golinkoff, 1986), they tailor their communication as mothers change their states of engagement (Lollis, 1987), they shift their gaze from the object to the addressee to supplement the force of their own messages (O'Neill, Astington, & Flavell, 1992) and they begin to give 'yes/no' answers to simple statements (Pea, 1980). Importantly, these observations are generally taken from felicitous, naturalistic, supportive and engaging environments with others; however, they do not show what infants' expectations are about the behavior of others. In a world in which language derives from vastly different kinds of sources, an examination of how infants attempt to make sense of labeling behavior across controlled contexts involving human and non-human labelers allowed us to examine infants' expectations regarding fundamentally different communicators.

The finding that infants respond differently to human and non-human labeling sources is consistent with other research indicating an ability in infants to make a distinction between behavior they expect from human, psychological agents and mechanical, inanimate objects. In research by Meltzoff and Moore (1995), 18-month-olds were presented with a person as well as a mechanical device responsible for three different failed attempts to accomplish a goal, i.e. to press a button with a stick or to pull apart two attached blocks. After infants witnessed the person make various failed attempts, they successfully completed the actions. In contrast, after watching a mechanical device operate on the blocks, infants did not respond based on the apparent goal. Furthermore, habituation research with 6- and 9-month-olds by Woodward (1998) found that infants selectively attended longer to a human agent's reach when it was extended towards a new goal object than when it was extended to the same goal across a new path. Interestingly, such results did not emerge when infants saw an inanimate entity with digits extend to grasp the toy. As early as 2–3 months of age, infants smile and vocalize more to responsive people than to interactive objects and they become distressed when people fail to communicate or respond contingently but not when inanimate objects fail to respond (Ellsworth, Muir, & Hains, 1993; Legerstee, 1997; Legerstee, Pomerleau, Malcuit, & Feider, 1987). By 6 months, infants expect people to behave differently to persons (i.e. by talking) than to inanimate objects (i.e. reaching and swiping; Legerstee, 2000). Our results extend such findings into the domain of linguistic behavior, showing that infants respond differently to animate vs. inanimate labeling sources. Our findings further complement the prior research in that infants looked to humans but not inanimate beings in the face of referential failures to seek information that may prove relevant to explanations about what went wrong.

There are a number of issues that current studies cannot address and that will require further research. As discussed, heightened looking to the true human labeler of Study 4

likely reflects an intentional understanding. Infants gave more attention to the true labeler of Study 4 than that of Study 1; such attention likely reflects their wonder at how a person could label accurately without visual access to the relevant information. It is also worth noting that infants' tendency to look longer to the true labeler than the false labeler of Study 4 was non-significant. Thus, infants' interest in both the true and false labelers of Study 4 may also be due to their interest in making the appropriate inference by determining the appropriate referent of a person's gaze. Because infants' looking to the back wall in the current study was considered a 'look-away' from the labeling event, infants' attention to the focus of a person's gaze in Study 4 would have been cut short. Furthermore, the position of infants in relation to the wall made it difficult for them to follow the speaker's gaze. A different set-up that made it easier for infants to follow the gaze of the turned-around labeler and that measured such attention would constitute steps toward distinguishing different possible interpretations infants may have.

Another issue concerns how one defines "familiarity" when conducting research on infants' comprehension of true and false labeling events. In the current set of studies, to guarantee that the False Condition indeed presented word–object violations to infants, both the words and the visual stimuli that were used were known, not novel, to the infant. In regard to the violated pairings of the False Condition (e.g. "ball" – [duck]), infants may recognize a violation both by seeing something different than what they expect when they hear "ball" and by hearing something different than what they expect when they see a duck. Because they know both what ducks and balls are, namely, what they look like and what they are called, current false word–object pairings are, in effect, two-sidedly false. Infants' ability to treat familiar true utterances as distinct from false utterances may hinge on the degree of familiarity between a word and its referent and the flexibility or rigidity of infants' conceptual repertoire. For example, given a bias of mutual exclusivity and infants' ability to link novel labels with novel objects (e.g. Golinkoff et al., 1995; Markman, 1989), an interesting question is whether a novel term, like "gip", and a familiar term, "chair", would be treated as equally unacceptable by infants as labels for [dog]. Indeed, in the real world, there will be cases in which infants must learn second names for an object (e.g. superordinate or subordinate labels such as "animal" or "poodle"). Such cases generally will involve novel rather than familiar terms. It is possible that the expectations that we have discussed, that is, expectations that words refer and expectations that people label truthfully, ultimately lead children to accept second names for objects after repeatedly hearing these names produced by a person.[3]

The current research was a first attempt to compare how infants treat two very different sources of information: words and speakers. It aimed to inform our understanding by addressing certain questions: when infants are presented with a mismatch between language and their experience, do they recognize the discrepancy between the claim and the world? If they do recognize these contradictions, do they, in any way, appreciate both of the critical sources of information, that is, the word and the speaker? The results of these four studies suggest that by 16 months of age, infants appreciate each source of information in light of their developing understanding of the relationship between words and the speakers who use them.

[3] We thank an anonymous reviewer for this observation.

Acknowledgements

Support for this research was provided by NICHD Grant HD30820 to the second author. A version of this report was submitted by the first author as a Master's Thesis to the University of Texas at Austin. We thank Catherine Caldwell, Don McGeary, and Hipolito Mata for their able research assistance and Cheryl Browne, Susan Gelman, Art Markman, Jacqueline Woolley and three anonymous reviewers for comments on earlier presentations of the work. We are grateful to the infants who participated, their families and Carmen Gonzalez-Sifuentes of the Children's Research Laboratory for their indispensable roles in this research.

Appendix A. Different orders of stimuli presentation in the True and False Conditions

Infants saw the items in brackets on the screen and heard the item in quotes from the labeling source.

A.1. True Condition

A.1.1. Order 1

1. [ball] "That's a ball"
2. [duck] "That's a duck"
3. [shoe] "That's a shoe"
4. [chair] "That's a chair"
5. [cat] "That's a cat"
6. [ball] "That's a ball"
7. [duck] "That's a duck"
8. [shoe] "That's a shoe"
9. [chair] "That's a chair"
10. [cat] "That's a cat"
11. [ball] "That's a ball"
12. [duck] "That's a duck"

A.1.2. Order 2

1. [cat] "That's a cat"
2. [chair] "That's a chair"
3. [shoe] "That's a shoe"
4. [duck] "That's a duck"
5. [ball] "That's a ball"
6. [cat] "That's a cat"
7. [chair] "That's a chair"

8. [shoe] "That's a shoe"
9. [duck] "That's a duck"
10. [ball] "That's a ball"
11. [cat] "That's a cat"
12. [chair] "That's a chair"

A.2. False Condition

A.2.1. Order 1

1. [shoe] "That's a ball"
2. [chair] "That's a cat"
3. [ball] "That's a duck"
4. [cat] "That's a shoe"
5. [duck] "That's a chair"
6. [shoe] "That's a duck"
7. [chair] "That's a cat"
8. [ball] "That's a shoe"
9. [cat] "That's a chair"
10. [duck] "That's a ball"
11. [chair] "That's a shoe"
12. [shoe] "That's a duck"

A.2.2. Order 2

1. [duck] "That's a chair"
2. [cat] "That's a shoe"
3. [ball] "That's a duck"
4. [chair] "That's a cat"
5. [shoe] "That's a ball"
6. [cat] "That's a duck"
7. [chair] "That's a shoe"
8. [duck] "That's a ball"
9. [shoe] "That's a cat"
10. [ball] "That's a chair"
11. [duck] "That's a shoe"
12. [cat] "That's a ball"

References

Astington, J. W., & Gopnik, A. (1991). Developing understanding of desire and intention. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 39–50). Cambridge, MA: Basil Blackwell.

Austin, J. L. (1975). *How to do things with words* (2nd ed.). Cambridge, MA: Harvard University Press. (Original work published 1962)

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

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

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

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

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

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mind-reading. *Cahiers de Psychologie Cognitive*, *13*, 513–552.

Barwise, J., & Perry, J. (1983). *Situations and attitudes*. Cambridge, MA: MIT Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, *21* (3), 205–226.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, *6*, 183–200.

Bratman, M. E. (1987). *Intention, plans and practical reason*. Cambridge, MA: Cambridge University Press.

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a "theory of mind"? In M. E. Lamb & L. R. Sherod (Eds.), *Infant social cognition* (pp. 333–373). Hillsdale, NJ: Erlbaum.

Burge, T. (1986). Individualism and psychology. *Philosophical Review*, *95*, 3–43.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, *15*, 317–355.

Clark, E. V. (1991). Acquisitional principles in lexical development. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: interrelations in development* (pp. 31–71). Cambridge: Cambridge University Press.

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, *97*, 256–266.

de Saussure, F. (1966). *Course in general linguistics* (pp. 65–78). New York: McGraw-Hill. (Original work published 1915)

Ellsworth, C., Muir, D., & Hains, S. (1993). Social competence and person-object differentiation: An analysis of the still-face effect. *Developmental Psychology*, *29*, 63–73.

Flavell, J. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. W. Astington, P. L. Harris & D. R. Olson (Eds.), *Developing theories of mind* (pp. 244–267). New York: Cambridge University Press.

Golinkoff, R. M. (1986). "I beg your pardon?": the preverbal negotiation of failed messages. *Journal of Child Language*, *13*, 455–476.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23–46.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1995). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, *21*, 125–155.

Gopnik, A., Slaughter, V., & Meltzoff, A. N. (1994). Changing your views: How understanding visual perception can lead to a new theory of the mind. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: origins and development* (pp. 157–181). Hillsdale, NJ: Erlbaum.

Grice, P. (1969). Utterer's meaning and intentions. *Philosophical Review*, *78*, 147–177.

Higginbotham, J. (1989). Knowledge of reference. In A. George (Ed.), *Reflections on Chomsky* (pp. 153–174). Oxford: Basil Blackwell.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*, Cambridge, MA: MIT Press.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E., & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish speaking infants and toddlers. *Journal of Child Language*, *20*, 523–549.

Kripke, S. (1972). *Naming and necessity*. Cambridge, MA: Harvard University Press.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*, 1138–1141.
Legerstee, M. (1997). Changes in social-conceptual development: domain-specific structures, self-organization, and indeterminism. In A. Fogel, M. C. D. P. Lyra & J. Valsiner (Eds.), *Dynamics and indeterminism in developmental and social processes* (pp. 245–260). Mahwah, NJ: Erlbaum.
Legerstee, M. (2000). Precursors to the development of intention at 6 months: Understanding people and their actions. *Developmental Psychology*, *36* (5), 627–634.
Legerstee, M., Pomerleau, A., Malcuit, G., & Feider, H. (1987). The development of infants' responses to people and a doll: Implications for research in communication. *Infant Behavior and Development*, *10*, 81–95.
Lollis, S. P. (1987). Effects of maternal behaviour on toddler preseparation and separation behaviour. *Dissertation Abstracts International*, *46*, 4037.
Markman, E. M. (1989). *Categorization and naming in children*. Cambridge, MA: MIT Press.
Markman, E. M., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cognitive Psychology*, *16*, 1–27.
Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*, 121–157.
Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838–850.
Meltzoff, A. N., Gopnik, A., & Repacholi, B. M. (1999). Toddlers' understanding of intentions, desires and emotions: explorations of the dark ages. In P. D. Zelazo, J. W. Astington & D. R. Olson (Eds.), *Developing theories of intention* (pp. 17–42). Mahwah, NJ: Erlbaum.
Meltzoff, A. N., & Moore, M. K. (1995). A theory of the role of imitation in the emergence of self. In P. Rochat (Ed.), *The self in early infancy: theory and research* (pp. 73–93). Amsterdam: North-Holland/Elsevier Science.
Merriman, W. B., & Bowman, L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development*, *54*(3–4, Serial No. 220).
Mervis, C. B., Golinkoff, R. M., & Bertrand, J. (1994). Two-year-olds readily learn multiple labels for the same basic level category. *Child Development*, *65*, 1163–1177.
Moore, C. (1999). Intentional relations and triadic interactions. In P. D. Zelazo, J. W. Astington & D. R. Olson (Eds.), *Developing theories of intention* (pp. 43–61). Mahwah, NJ: Erlbaum.
Naigles, L. G., & Gelman, S. A. (1995). Overextensions in comprehension and production revisited: Preferential-looking in a study of *dog*, *cat*, and *cow**. *Journal of Child Language*, *22*, 19–46.
Nelson, K. (1973). Structure and strategy in learning to talk, *Monographs of the Society for Research in Child Development*, *38*(1–2, Serial No. 149).
O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, *67*, 659–677.
O'Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development*, *63*, 474–490.
Oviatt, S. L. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development*, *51*, 97–106.
Partee, B. (1982). Belief sentences and the limits of semantics. In S. Peters & E. Saarinen (Eds.), *Processes, beliefs, and questions* (pp. 87–106). Dordrecht: D. Reidel.
Pea, R. D. (1980). The development of negation in early child language. In D. R. Olson (Ed.), *The social foundations of language and thought: essays in honor of Jerome S. Bruner* (pp. 156–186). New York: W.W. Norton.
Perner, J. (1991). *Understanding the representational mind*. London: MIT Press.
Putnam, H. (1975). The meaning of meaning. In K. Gunderson (Ed.), *Language, mind and knowledge*. Minneapolis, MN: University of Minnesota Press.
Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, *69*, 309–320.
Searle, J. (1969). *Speech acts*. Cambridge: Cambridge University Press.
Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*, New York: Cambridge University Press.
Segal, G. (1989). A preference for sense and reference. *Journal of Philosophy*, *86*, 73–89.
Shatz, M., & O'Reilly, A. W. (1990). Conversational or communicative skill? A reassessment of two-year-old children's behavior in miscommunication episodes. *Journal of Child Language*, *17*, 131–146.

Spelke, E. S., Phillips, A., & Woodward, A. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack & A. J. Premack (Eds.), *Causal cognition: a multidisciplinary debate* (pp. 44–78). New York: Oxford University Press.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal correspondences by human infants. *Science*, *222*, 179–181.

Swingley, D., & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language*, *46*, 39–56.

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

Tomasello, M., Farrar, M.J., & Dines, J. (1984). Children's speech revisions for a familiar and unfamiliar adult. *Journal of Speech and Hearing Research*, *27*, 359–363.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press/Bradford Books.

Wellman, H. M. (1993). Early understanding of mind: The normal case. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds: perspectives from autism* (pp. 10–39). New York: Oxford University Press.

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

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 65–77.

Woodward, A. L., & Markman, E. M. (1998). Early word learning. In W. Damon (Series Ed.) & D. Kuhn, & R. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception and language* (5th ed., pp. 371–420). New York: Wiley.