THIRD PARTY SOCIAL INTERACTION AND WORD LEARNING FROM VIDEO

By

Katherine D. O'Doherty

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

In

Psychology

August, 2009

Nashville, TN

Approved By:

Georgene L. Troseth

Megan M. Saylor

ACKNOWLEDGEMENTS

This research would not have been possible without the financial support of Peabody College at Vanderbilt University and the Learning Sciences Institute.

I would like to thank Dr. Megan Saylor at Vanderbilt University for being on my committee and for her helpful contributions. Also, many thanks to Gabrielle Strouse, Brian Verdine and the Vanderbilt Early Development Lab research assistants, especially Lauren Deisenroth, Paige Holden and Alison Milam for their help with data collection and coding. I am also very thankful to my advisor, Dr. Georgene Troseth for her continuous guidance, encouragement, and helpful contributions to all aspects of this project.

I would also like to thank Dr. Nameera Akhtar, Dr. Priya Shimpi and Elizabeth Goldenberg at University of California, Santa Cruz for their help with planning, data collection and insights throughout the research process.

Finally, I would like to thank my family and my husband, Ryan.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS ii

LIST OF FIGURES iv

Chapter

I. INTRODUCTION 1

II. STUDY 1 7

Method 7
Participants 7
Materials 8
Procedure 9
Coding 13
Results 13
Discussion 16

III. STUDY 2 17

Method 17
Participants 17
Materials and Procedure 17
Results 18
Discussion 19

IV. GENERAL DISCUSSION 20

V. REFERENCES 29

LIST OF FIGURES

Figure Page

1. Novel test objects 8

2. Views of the (a) researcher in the *Live* and *Video Addressed* conditions and the (b) researcher and confederate in the *Live* and *Video Onlooker* conditions. 11

3. Percentage of children who learned the novel word 15

4. Percentage of children who learned the novel word (Studies 1 and 2) 18

CHAPTER I

INTRODUCTION

From early infancy, children are exposed to information from varied sources, including parents and siblings, overheard conversations ("third party" social interaction - Akhtar, 2005), and television. A general challenge young children face is determining when and from whom to take information. By age two, children are sensitive to the presence or absence of referential social cues, and thus are quite skilled at recognizing what constitutes a learning opportunity. Baldwin and Moses (2001) relay an anecdote in which their 2-year-old son, playing with new toys while they watched the evening news, suddenly declared, "No legal precedent!" Although he obviously had liked the sound of the phrase when it had emanated from the TV, this discerning youngster realized that the words did not refer to the objects he was examining, and did not begin to use them as object labels. The authors point out the frequency of errors that would occur if children relied on temporal associations to learn words, rather than on social "clues" to a speaker's referential intentions.

Across the first year of life, children become aware of a range of such "clues" that indicate the intention to communicate information. Csibra and Gergely (2006) argue that from birth, infants are aware of social signals including eye contact and the prosody of infant-directed speech. Young infants' attention is directed by such cues: newborns attend more to an individual using infant-directed than adult-directed speech (Cooper & Aslin, 1990) and 6-month-olds follow the gaze of a person who first makes eye contact with

them (but not a person with averted gaze) to a specific target object (Csibra & Gergely, 2009). Later in the first year, according to these theorists, social signals indicate to infants that a "pedagogical" or teaching situation is taking place, facilitating knowledge transfer. Throughout the first years of life, the ability to read meaning into social signals becomes increasingly sophisticated. By 9 to 12 months, infants use cues such as eye gaze to draw conclusions about an actor's goals (Woodward, 2003) and, shortly after their first birthday, to share an understanding of an actor's intentions (Tomasello, Carpenter, Call, Behne & Moll, 2005).

By their second year, children are skilled at recognizing opportunities to learn from other people. Individual social cues (e.g., infant-directed prosody) do not appear to automatically "trigger" toddlers' perception of a learning situation; rather, children rely on evidence from the context in which these cues are presented to determine when intentional communication is occurring. For example, 18-to 20-month-olds learned a novel object label uttered by a person looking at the object with them; in contrast, an utterance in infant-directed speech from a disembodied voice, contingent on children's own attention to an object but lacking cues to reference, was not sufficient for children's learning (Baldwin, Markman, Bill, Desjardins, Irwin, & Tidball, 1996). In the context of social interaction, when a particular cue (e.g., information about gaze direction) is uninformative, toddlers adaptively use any of a variety of communicative cues (such as emotional expressions) that may be available (Akhtar & Tomasello, 1996; Tomasello & Barton, 1994; Tomasello, Strosberg, & Akhtar, 1996).

Along with their awareness of social cues, toddlers appear to understand, and can flexibly adopt, the complementary roles involved in social interaction. In one study, an

adult taught 18-month-olds a "placing" game (held out a plate for the child to deposit a toy); when given the plate, the children took the adult's role, offering the plate to allow the adult to place the toy (Carpenter, Tomasello, & Striano, 2005). Carpenter and her colleagues point out that collaborative role reversal such as this involves recognizing and then adopting the intentions of the other, and seeing the self and other as in some ways "the same" and interchangeable.

By the middle of the second year, children also appear to have developed expectations about the reliability of information communicated by social others that they do not extend to non-social sources. Sixteen-month-old children looked longer when an adult seated next to them *looked at and mislabeled* an object (called a shoe "*a ball*") compared to when the person correctly labeled the object (Koenig & Echols, 2003). However, they looked longer at an adult whose back was turned to an object who *correctly labeled* it (compared to when labeling was incorrect). Children of this age apparently expected accuracy from an adult who offered referential social cues, but not when such cues were missing. Children's attention to an audio speaker did not differ whether a correct or an incorrect object label came out of it, suggesting that children had no expectations of correct reference from a human voice in the absence of a person providing referential cues.

Reliance on the presence of others and the cues they provide may at least partially explain why very young children seem not to learn efficiently from video, a source of information that is becoming increasingly pervasive in infants' and toddlers' lives (Rideout & Hamel, 2006). This result, dubbed the "video deficit" (Anderson & Pempek, 2005) has been found at several ages and across various learning tasks. For example, at

an age when infants' speech perception is narrowing toward their parents' language, 9-month-olds maintained the ability to discriminate non-native (Mandarin Chinese) speech sounds after a Mandarin speaker interacted with them face-to-face for 5 hours across the course of a month, but other 9-month-olds lost this ability after watching the same speaker on video for the same amount of time (Kuhl, Tsao & Liu, 2003). In other research, 12- to 30-month-olds imitated the actions of a person who was present more often than they imitated the same person appearing on a TV screen, even though she made apparent eye contact and offered attention-directing comments (e.g., "Look at this!") in both cases (Barr & Hayne, 1999; Hayne, Herbert & Simcock, 2003). Similarly, 24-month-olds followed the instructions of a person who was present, using this information to find a hidden toy, but were only a third as likely to use the same information offered by the same person on video (Schmitt & Anderson, 2002; Troseth, Saylor & Archer, 2006). Although 30-month-olds (compared to younger children) showed some improvement in using information from people on video in search tasks (Schmitt & Anderson, 2002; Troseth & DeLoache, 1998), children of this age still did better when instructed by a person who was present (Schmitt & Anderson, 2002). In a difficult labeling task that demanded reliance on referential social cues, 24- and 30-month-olds learned a word when an adult who was present gazed into an opaque container while offering the novel label, ignoring a visible distracter; in contrast, toddlers of the same age did not learn the word from a person on video who offered the same cues (Troseth, Saylor, & Strouse, 2009; also see Krcmar, Grela & Lin, 2007).

In all of these studies, a person on video offered particular social cues, such as apparent eye contact with the viewing child and infant-directed language. However, other

aspects of a normal social situation were missing. For instance, the person on television did not engage in contingent, reciprocal interaction with the child—a characteristic of social engagement to which infants are sensitive from the middle of the first year (Bigelow, MacLean, & MacDonald, 1996; Hains & Muir, 1996). Additionally, the person on screen did not share attention with the viewing child to objects present in the child's environment, as usually occurs in a triadic interaction. TV watching may not have met toddlers' expectations for a social, "pedagogical" situation in which information relevant to the child is being transmitted. An exception is a study in which a person on video provided evidence of engagement and relevance by conversing with the child's parent via 2-way closed-circuit video while the child watched. The person on TV then played "Simon Says" with the child, talked about and discussed an item *in the child's environment* (a sticker on the child's shirt), and responded contingently to whatever the child and parent said and did. After interacting with the person on video for 5 minutes, 24-month-olds learned efficiently from her (Troseth et al., 2006; also see Nielsen, Simcock, & Jenkins, 2008).

Do children need to be involved in a contingent, reciprocal interaction with another person in order to learn from that person? Young children also learn as *onlookers* to third-party interactions (Akhtar, 2005). Children as young as 18 months used an adult's referential behavior to learn words while "overhearing" an exchange between the speaker and another person who also was present in the room, learning as well as children who were directly addressed by the speaker (Akhtar, Jipson & Callanan, 2001; Floor & Akhtar, 2006). It is important to note that the speaker's cues were not directed to the "overhearing" child; the speaker made eye contact and interacted only with the adult

confederate while stating her intention to show that person a named object. She removed the item from a container, held it up with a gasp of pleasure, demonstrated its function, and handed it over to the other adult before returning it to the container. No social behavior was directed at the children, who nevertheless learned the word as easily as children who were directly addressed.

CHAPTER II

STUDY 1

In the current study, we examined whether 30-month-olds learn from a person on video if expected social cues are offered, *but are directed toward another adult on the screen*. We predicted that children would learn from watching an interaction between two adults on television but would not learn from being directly addressed by a non-contingent person on television. Based on previous research, we expected that children would learn from someone present in their environment whether they were onlookers to a conversation or directly addressed. We used the original procedure of Akhtar and her colleagues (2001), with modifications that were required to equate live and video versions of the procedure.

Method

Participants

Sixty-four children (31 males) participated, ranging in age from 27.0 to 31.7 months, $M = 29.8$ $SD = 0.9$. Half of the participants (15 male), from a city in the southeast United States, were recruited from state birth records. The rest of the participants (16 male) from a community on the U. S. west coast, were recruited from a database of children whose parents had expressed interest in being included in studies of child development. In both of the studies reported here, participants were mostly of

European-American descent and were native English speakers. Across both studies, primary caregiver education ranged from high school diploma to a doctoral/professional degree (85% had a college degree or above) and family income from $30,000 to over $100,000 a year (79% earned $50,000 or above).

Children were randomly assigned to one of four conditions: *Live Addressed* (*M* = 29.7 months, *SD* = .80, 7 boys) *Live Onlooker* (*M* = 30.1 months, *SD* = .80, 8 boys) *Video Addressed* (*M* = 29.8 months, *SD* = 1.0, 8 boys), and *Video Onlooker* (*M* = 29.7 months, *SD* = .80, 8 boys). Data from 12 children were excluded from analysis for uncooperativeness (6), English not being the child's primary language (2), parental interference (1), suspected developmental delay (1) and experimenter error (2).

Materials

Four familiar items (e.g., plastic horse, banana, turtle, and truck) were used for a warm up comprehension task. Four distinctive wooden toys with movable parts were used as novel test objects (see Figure 1). All of these objects afforded interesting actions that 2-year-olds could perform.

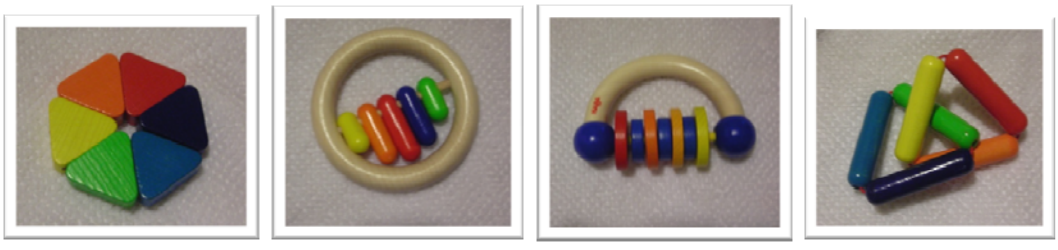

Figure 1: Novel test objects

Four different-colored opaque jars with screw-on lids, placed in a row, and attached to a 36” (91 cm) wood plank were used as the hiding apparatus for the objects. During the warm-up and testing phases, participants sat at a small table across from an experimenter while parents sat on a nearby couch. During the labeling of the novel objects, children sat on the couch or on a small chair, facing the researcher(s) in the live conditions and a 27” or 32” (69 or 81 cm, depending on location) television set that displayed a pre-recorded video of the researcher(s) in the video conditions. One video camera filmed the entire experimental setup, including child and researchers/TV. A second camera was directed at the participant during the labeling phase; the resulting video was used to code children’s attentiveness.

Procedure

Testing in both locations took place in small laboratory playrooms and lasted approximately 30 minutes. The researcher explained the study to parents, obtained consent, and then asked parents to complete the short form of the MacArthur Communicative Development Inventories (Level II) to assess their child’s expressive vocabulary (Fenson, Pethick, Renda, Cox, Dale & Reznick, 2000). The researcher also confirmed that none of the children were familiar with any of the four novel objects.

Warm up. After a brief warm up activity, the researcher introduced a showing game and comprehension test using the four familiar items. After placing the hiding apparatus on the table, the researcher said, “I’m going to show you what’s in here”, removed the lid from the first container, and pulled out the familiar item inside. She then handed the item to the child, allowed the child to handle it briefly, replaced this object in

the container and continued until all four containers had been opened and their contents examined. After introducing the four familiar items, the researcher placed all the items in a tray in front of the participant and asked the child to choose one (e.g., "Can you pick the horse?"). After the child correctly chose the requested item, the experimenter replaced it and asked the child to choose a different object, until the child had correctly picked two items in a row, indicating they understood the instructions.

Next, the researcher placed the hiding apparatus, now filled with the four novel objects, on the table and said, "I'm going to show you what's in here. Let's see what's in here. I'm going to show you this one". She then removed the lid from the first jar, pulled out the novel item inside, handed the item to the child and then asked the child to return it to the jar. These steps were repeated for the remaining containers until the child had been familiarized with all four novel objects. No labeling occurred during this warm up. The warm-up familiarized the children with the novel objects and the apparatus, and confirmed that they understood what they were being asked to do in the comprehension task.

Labeling. In the *live* conditions, the child sat on the couch approximately 3-4 feet (1 meter) from the researcher who was directly facing the child (*Live Addressed*) or an adult confederate (*Live Onlooker*). In the *video* conditions, the child sat approximately the same distance from the television set which showed a pre-recorded video of the researcher facing the camera (*Video Addressed*) or an adult confederate (*Video Onlooker*) (see Figure 2).

a.)  b.) 

Figure 2: Views of the (a) researcher in the *Live* and *Video Addressed* conditions and the (b) researcher and confederate in the *Live* and *Video Onlooker* conditions.

In all four conditions, the researcher sat at a small table that held the containers. Each of the novel objects was always placed in a particular container, but which object was the target was counterbalanced across children. Prior to introducing the *target object* the researcher stated its label three times: "I'm going to show you the toma. Let's see the toma. I'm going to find the toma". Prior to introducing each of the distracter objects the researcher made three non-labeling statements: "I'm going to show you this one. Let's see this one. I'm going to find this one".

In the *Live Addressed* and *Video Addressed* conditions, all words and actions were directed at the child: the researcher made eye contact with the child (or the camera) and uttered the three statements, then opened the container, looked inside while gasping and smiling, pulled out the object, performed a distinctive action on it (e.g., shaking it; rocking it from side to side) for approximately five seconds, replaced the object in the container and, moved on to the next one.

In the *Live Onlooker* and *Video Onlooker* conditions, the same words and actions were directed at the confederate instead of the child. The speaker performed the distinctive action on the object for approximately two seconds, and then handed the

object to the confederate who imitated the action for approximately three seconds before handing the object back to the speaker. The *Live Onlooker* condition was the same as that used previously by Akhtar et al. (2001). The *Live Addressed* condition was the same as Akhtar et al.'s *Addressed* condition, except that the researcher did not hand the toys to the child (this was done to equate the *Live* and *Video* conditions). In all conditions, each object was out of its container and visible to the child for approximately five seconds. In no condition did the child touch the toys during the labeling phase.

In all conditions, the researcher proceeded from right to left through each of the four containers in the apparatus, repeating the process for a total of three times. Thus, children heard the target object labeled the "toma" a total of nine times. This demonstration took approximately three minutes. Once labeling began, the researcher/video was not stopped for any reason (e.g., if the child got up or turned away from the demonstration).

Testing. Children sat across from the researcher at the small table and were allowed to handle the novel objects (without labeling) for approximately 30 seconds. The objects were then put into a tray, which the researcher shook up and placed in front of the child. The researcher looked up from the tray, made eye contact with the child, and asked a comprehension question, "Which one is the toma?" as well as a preference question, "Which one is your favorite?". The order of questions was counterbalanced across children. To ensure that children did not think that the two questions were the same, there was a minimum of 30 seconds between questions in which the researcher asked about the child's outfit, siblings, etc. If the child did not pick an item or picked several, the question was repeated until the child chose one object for each question.

Coding

Attention. As a measure of attentiveness, the proportion of time children spent looking at the researcher(s) or video during the labeling session was independently recorded for all of the videotaped sessions by one coder. An additional coder recorded children's attentiveness for 25% of the videotaped sessions (four participants from each condition). Inter-rater reliability was high ($\rho_I = .96$, $p < .001$). An additional coder, blind to hypotheses, also coded 25% of the videotaped sessions; reliability with the main coder was exactly the same.

Learning. An assistant recorded the participant's answer to the comprehension and preference questions during the session. We considered children who chose the target in response to the comprehension question, but not in response to the preference question, to have demonstrated learning. An independent coder who was blind to the hypotheses viewed all of the videotaped sessions and recorded whether or not the child had learned; agreement between the two raters was 95% ($\kappa = .93$, $p < .001$).

Results

Vocabulary. In preliminary analyses, we compared children's raw scores on the MacArthur CDI. Girls' scores ($M = 90.2$, $SD = 16.9$) were significantly higher than boys' ($M = 79.1$, $SD = 20.8$); $t\ (62) = 2.35$, $p = .02$, $\eta^2 = .08$), but there were no significant differences by age ($t\ (62) = .34$, $p = .74$, $\eta^2 = .002$) or testing location ($t\ (62) = -.94$, $p = .36$, $\eta^2 = .01$). In a one-way analysis of variance, there were no significant differences in vocabulary scores across the four conditions ($F\ (3,60) = .05$, $p = .99$, $\eta^2 = .002$).

Attention. In preliminary tests of children's attentiveness, there were no effects of gender ($t(60) = -1.18$, $p = .24$, $\eta^2 = .02$), age ($t(60) = .83$, $p = .41$, $\eta^2 = .01$) or testing location ($t(60) = -.37$, $p = .71$, $\eta^2 = .002$). A two-way between groups analysis of variance exploring the effects of medium (Live vs. Video) and conversation type (Addressed vs. Onlooker) on children's attentiveness revealed a significant a main effect of medium ($F(1, 58) = 6.12$, $p < .05$, *partial* $\eta^2 = .095$). Children in the *Live* conditions spent a higher proportion of time ($M = 92\%$, $SD = .07$) looking at the demonstration than children in the *Video* conditions *(*$M = 84\%$, $SD = .15$*).* There was no main effect of conversation type ($F(1,58) = .38$, $p = .54$, *partial* $\eta^2 = .006$): children's attentiveness in the *Onlooker* conditions ($M = 87\%$, $SD = .14$) and the *Addressed* conditions ($M = 89\%$, $SD = .11$) was nearly equivalent. There was no interaction between medium and conversation type ($F(1,58) = 1.59$, $p = .21$, *partial* $\eta^2 = .027$). Two participants were excluded from attention analyses because of technical problems with their videotapes.

Learning. In preliminary tests, there were no effects of age ($\chi^2(1, N = 64) = .003$, $p = .96$, *phi* = -.01); or testing location ($\chi^2(1, N = 64) = .59$, $p = .44$, *phi* = .10); on children's learning. As is typical in tests of early word learning, girls performed significantly better than boys ($\chi^2(1, N = 64) = 4.44$, $p = .04$, *phi* = .26); however, because the number of girls and boys was balanced for each condition, this gender difference did not affect our analysis of differences in learning across conditions.

For each condition, we ran a binomial test to compare the number of children who learned the novel word against the number expected by chance (based on choosing the one target out of four objects for the *comprehension* question and choosing from the three non-target objects for the *preference* question: 0.25*0.75 = 0.1875). Both *Onlooker*

conditions *learned* at rates that exceeded chance (*Live Onlooker*, 9 of 16 children and *Video Onlooker*, 8 of 16 children, *ps* = .001 and .005, respectively). In contrast, in the two *Addressed* conditions (*Live* and *Video*), only 4 of 16 participants chose the target for comprehension but not preference, rates that were not significantly above chance (both *p*s = .35) (See Figure 3).

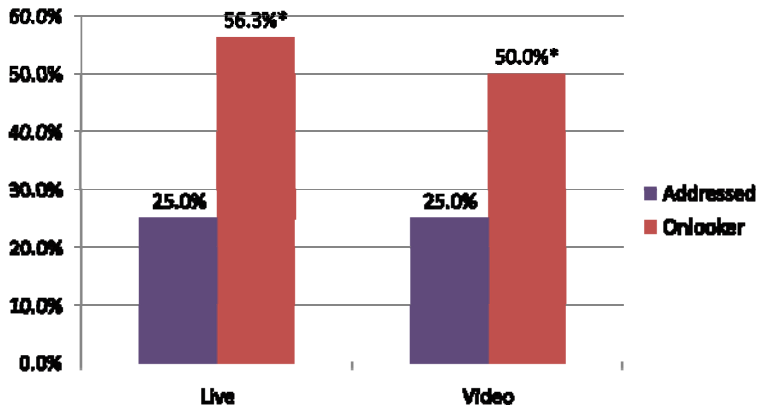

Figure 3: Percentage of children who learned the novel word

Children in the *Onlooker* conditions performed significantly better than children in the *Addressed* conditions (χ^2 (1, N = 64) = 5.32, p = .021, *phi* = -.3). Children's performance in the *Live* conditions did not significantly differ from children's performance in the *Video* conditions (χ^2 (1, N = 64) = .07, p = .80, *phi* = -.03). Therefore, children who were directly addressed by a person who was present or on video did not demonstrate learning of the novel word. In contrast, children who observed the person addressing another adult and labeling the novel object (either "in person" or on video) did

learn the word. Note that although attention was higher for the live events than for the video events, this did not predict the pattern of learning.

We also analyzed children's answers to the comprehension question without excluding those who also preferred the target. The patterns were the same, except the chi-square analysis for conversation type became non-significant (χ^2 (1, N = 64) = 3.07, p = .08, *phi* = -.22).

Discussion

Based on previous research (e.g., Akhtar, 2005; Akhtar et al., 2001) we expected children in the *Live Addressed* condition to learn the novel label. In Akhtar's previous studies, the researcher handed the children in the *Addressed* condition each object after talking about it (in the *Onlooker* condition, she handed the object to the adult confederate). In the present study, in order to match our *Video Addressed* condition, the speaker in the *Live Addressed* condition did not hand the child the toy after discussing it. We assume that a triadic interaction between a toddler and adult regarding new toys—especially one that involved multiple comments on and displays of each object—typically would include offering the child the objects of interest. In addition, the researcher had offered the child the toys when she was introducing the apparatus and finding game during the warm-up. In our *Live Addressed* condition, we may have violated toddlers' expectations of what a social *interaction* should include. Although handling the toys immediately after labeling was not necessary for learning (the children in the Overhearing conditions were not handed the toys), the presence of an obviously engaged social partner may have been vital.

CHAPTER III

STUDY 2

We hypothesize that unfulfilled expectations regarding normal social interaction hindered children's learning. To test this hypothesis, in Study 2, a group of children participated in a *Modified Live Addressed* condition in which children were handed each object after labeling and thus clearly were included as part of the interaction.

Method

Participants

Sixteen children (8 male) participated, drawn equally from the same two communities and populations, and recruited in the same way as in Study 1. Participants ranged in age from 28.5 to 31.6 months (M = 30.2; SD = .82).

Materials and Procedure

The same materials were used as in Study 1. The procedure was the same as in Study 1, *Live Addressed* condition, with some slight variations. During the *Labeling* phase, the researcher and child sat across the small table from each other (slightly closer than in Study 1, to allow the toys to be passed from researcher to child). Different from Study 1, after the speaker pulled each object from the container and performed an action on it, she gave the object to the child to handle briefly before replacing the object. Children played with the object for an average of 5 seconds (*range* = 0 to 17 seconds).

All children handled most of the objects; there were no participants that did not want to handle any of the objects or were unwilling to give the object back to the speaker at the appropriate time.

Results

In preliminary tests, we checked whether the length of time that children handled the target object (M = 5 seconds, range = 0 to 17 seconds) was related to learning the novel word using Spearman rank order correlation. There was no significant correlation between target handling time and learning (ρ = -.17, p = .54).

We used a binomial test to compare the number of children who learned the novel word against the number expected by chance (chance = 0.1875, as in Study 1). Nine of 16 participants learned the word, a number that is significantly above chance (p = .001) (See Figure 4).

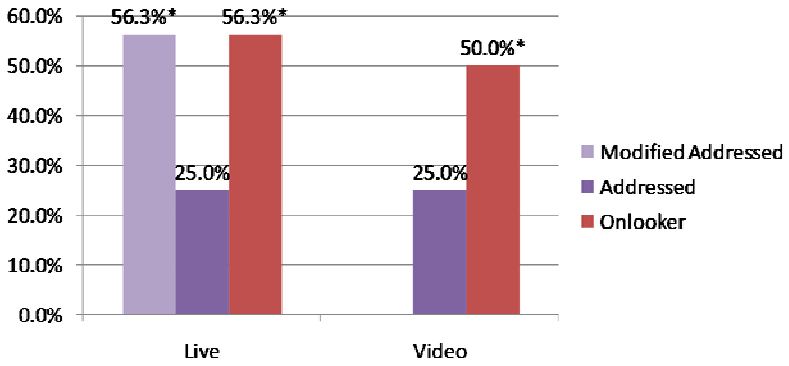

Figure 4: Percentage of children who learned the novel word (studies 1 and 2)

Therefore, when an adult addressed children and then handed the toys to them, they learned the novel word. The difference in children's learning in the *Modified Live Addressed* condition, compared to that in the original *Live Addressed* condition, approached significance (χ^2 (1, N = 32) = 3.24, p = .072, *phi* = .32).

Discussion

When we incorporated handing the object to children into our original *Live Addressed* procedure, children learned. During the warm-up to the showing game, children in all conditions were handed the novel objects, which may have raised their expectations that this was how the game was played. Sharing objects clearly indicates a reciprocal engagement with the social partner (whether one directly benefits from such sharing, as occurred in Study 2, or merely observes it, as in the *Onlooker* conditions of Study 1). Without evidence of reciprocal engagement, toddlers may not have oriented to the pedagogical cues and may not have recognized that the speaker had shared information.

CHAPTER IV

GENERAL DISCUSSION

The results of Study 1 indicate that 30-month olds can learn a novel word as an onlooker to a conversation between two people on video as well as they learn from a live third-party conversation in their environment. This study expands on previous research by providing evidence of a learning situation involving video in which toddlers do not contingently interact with the person on screen, yet learn as well as they do from overhearing those who are actually present. Consistent with previous research, we found that toddlers did not reliably learn a new word after being directly addressed by a person on a pre-taped video. An unexpected result in Study 1 was that toddlers did not reliably learn a new word after being directly addressed by a person who was present, either. Results from Study 2 indicate that this unexpected finding may be due to the live interaction not being truly reciprocal: children did learn the word when directly addressed by someone present in their environment who made the interaction reciprocal by handing them the objects.

In the current study, toddlers learned a novel name for an object as an onlooker to a conversation on video, as an onlooker to a conversation present in their environment, and when directly addressed by someone in their environment who offered them the objects of interest. What signaled to children that each of these scenarios was a learning situation? We believe the key element of these three contexts is *reciprocal interaction*. That is, children in these situations were either onlookers to knowledge exchanged via

social interaction or actively took part in a reciprocal interaction. Specifically, in the two overhearing conditions, the confederate watched intently as the speaker talked about the object to be revealed, then followed the speaker's gaze as she looked into each container and extracted an object. The speaker held each object up and manipulated it, then handed it to the confederate who briefly manipulated the object in the same way. This scene may have visibly demonstrated to onlooking children the occurrence of teaching (and learning). In the *Modified Live Addressed* condition (Study 2), the speaker handed each object to the viewing child after manipulating it, thereby allowing the child to participate in a reciprocal interaction. In the two conditions in which children did not learn (*Video* and *Live Addressed*) children did not take part in or observe a reciprocal interaction; they were mere observers of an extended, one-sided offering of information.

By 30 months of age, children are becoming discerning consumers of socially transmitted knowledge. We believe toddlers are sensitive to typical components of social "scripts" (both as observers and as participants), and in the absence of expected components may fail to recognize a learning opportunity. We hypothesize that toddlers are especially attuned to *reciprocal social interaction*, including behaviors exhibited by both actor and recipient in social games (Carpenter et al., 2005). By 12 to 15 months of age, infants have begun to take an active role in collaborative learning (Tomasello et al., 2005). For instance, in a study examining 15- and 18-month-old infants' behavior during adult-initiated interruptions in social game playing, 60% of infants' actions were interpreted as communicative attempts to reengage the adult in the game (Ross & Lollis, 1987). In addition to taking an active role in social learning, older infants apparently count on others to play their role in a typical manner (displaying expected social cues).

For example, 14-month-olds did not imitate a demonstrated action in an "incidental learning context" where the actor did not provide any communicative cues (Király, Csibra & Gergely, 2004). Nielsen (2006) found that 18-month-olds were significantly more likely to imitate an adult's exact actions when she demonstrated engagement (i.e., played with child in the warm up, made eye contact and smiled during the demonstration, alternating gaze between toy and child) than when she acted aloof (no prior playing, no eye contact, no smiling), whereas 24-month-olds were equally likely to imitate regardless of the adult's demeanor. The author hypothesized that in imitating the person who was engaged with them, children of both ages showed they wanted to sustain interaction; additionally, the older children may have been attempting to *initiate* interaction by imitating the aloof actor. Toddlers' strong motivation to attend to and engage in reciprocal interactions therefore may facilitate their learning.

In the *Addressed (Live* and *Video)* conditions of Study 1, following a warm-up in which children were allowed to touch each item after it was removed from its container, children were placed in the role of observer of a one-sided labeling demonstration that went on for 3 minutes. Although the adult carrying out the demonstration made eye contact with and directed remarks to children (following a script), she may not have seemed fully engaged with them, given their prior experience in the showing game during warm up. Procedural differences may explain why in other research (e.g., studies of early imitation) infants and toddlers have learned after merely observing an individual's behavior with a novel toy (e.g., Barr, Dowden, & Hayne, 1996; Meltzoff, 1985). As part of the imitation procedure, children typically are not given access to the toy before the novel behaviors are modeled, and demonstrations are very short (20 to 60 seconds). In

another type of study, infants and toddlers viewed an adult's communicative cues and learned the location of a hidden toy (Behne, Carpenter, & Tomasello, 2005). However, children did not need to wait an extended period of time for "their turn" to find the toy, which immediately followed a very brief presentation. In contrast, the researcher in the current studies discussed, revealed, and manipulated each of the four objects three times; thus, in the *Addressed* conditions in Study 1, children sat through 12 object "showings" before it was their turn to play. Toddlers in the *Modified Live Addressed* condition, who were given the opportunity to play with each toy right away (as had occurred in the warm-up), learned the novel word. Learning was not related to how long children handled the target object, indicating that the *handing over* of the toys (i.e., the evidence of the researcher's engagement) was the crucial part.

Akhtar and colleagues' research (and the research reported here) indicates that toddlers do not need to be an active part of a social interaction to view it as a learning situation; they learn as an onlooker to a conversation as well as when they are directly addressed. In Akhtar's paradigm (and our own) the recipient in the *Onlooker* learning situation is an important part of the interaction; she handles each novel object and demonstrates that she has learned to perform an action on it. In a recent study, Herold and Akhtar (2008) hypothesized that the ability to learn as an onlooker relies on understanding "self-other equivalence" (Moore, 2007), because it allows children to imagine themselves as part of the interaction. They assessed 18- to 20-month-olds' self-recognition and their ability to take another person's perspective. Both factors predicted children's imitation of novel behaviors that the children learned as onlookers to a "third-party interaction". We hypothesize that children learn as onlookers (both in this study and

from the real interactions they observe daily) because they recognize that knowledge is being transferred via social interaction (and possibly imagine themselves as part of the exchange). A question for future research is whether or not toddlers would continue to learn as onlookers to a situation in which the knowledge recipient (the confederate) is an observer of a one-sided demonstration. That is, if the recipient does not provide any sign of being part of a reciprocal interaction, would toddlers still learn the information being presented?

In the current study, toddlers exhibited no "video deficit" in learning after observing a social interaction on video, in contrast to many other studies in which learning from video was depressed compared to learning from an equivalent "live" event (e.g., Barr & Hayne, 1999; Deocampo & Hudson, 2005; Hayne, et al., 2003; Kuhl, et al., 2003; Schmitt & Anderson, 2002; Strouse & Troseth, 2008; Troseth et al., 2006). In a recent word learning study, 24- and 30-month-olds failed to use subtle referential social cues (e.g., the labeler's gaze into an opaque bucket containing a target object, in the presence of a visible distracter) to learn a word from a person on a video compared to a person who was present (Troseth, et al., 2009). Even when a person's social cues were straightforward (gaze toward one of two visible toys), toddlers more often learned a word on trials when a person was present compared to on video trials (Krcmar et al., 2007). Additionally, 30- to 35-month-old children did not learn verbs after repeatedly watching an event on video narrated in a voiceover using infant-directed speech; they did learn when the first two video demonstrations were replaced with live social interaction (an adult using a doll or puppet to demonstrate and label the action—Roseberry, Hirsh-Pasek,

Parish-Morris, & Golinkoff, in press). Only children over age 3 learned the verbs from video alone.

The presence of an engaged recipient of the social interaction, as in the current research, may have promoted toddlers' awareness of the relevance of referential social cues presented on video. Kuhl (2007; Kuhl et al., 2003) points out that a basis for infants' language acquisition may be seeing a person's social cues as referential; social cues that are informative when the speaker is present in the environment and directing attention to real objects may not seem referential coming from a non-contingent person on video directing his or her gaze/points/actions at 2-dimensional images of objects on the screen. In the present study, when both parties to the interaction were together on screen and an obvious exchange of information took place, toddlers appeared to treat this scene as an opportunity to learn.

Nielsen et al. (2008) suggest ways in which social interaction affects learning from video. In their study, a modeler demonstrated an arbitrary and somewhat awkward action (using a stick to press a switch to open a box, rather than using hands). Toddlers were more likely to imitate the exact behaviors of a person who was present and responsive than a person on a pre-taped video. In a follow-up study, children were more likely to imitate the exact actions of a modeler who had been contingently responsive to them via closed-circuit video; in contrast, when the same modeler on a pre-taped video was non-responsive, toddlers tended to open the box with their hands. Nielsen et al. reason that social interaction affected imitation because children viewed the responsive person on video as a social partner with whom they could affiliate (forge an interpersonal bond and sustain interaction—Meltzoff & Moore, 2002; Uzgiris, 1981). In the current

study, toddlers observed an interaction on video between a teacher and a learner who copied the teacher's exact actions, sustaining the interaction. The teacher treated the learner as she had the child during the warm-up, handing over each toy to her. Viewing children may have learned by "putting themselves in the place of" the responsive learner on the screen (Herold & Akhtar, 2008).

Referential cues help children determine the intended target in situations where the object is not visible during labeling (as in the current studies) or when there is more than one possible referent for an utterance (Baldwin, 1993; Baldwin & Moses, 1996). For toddlers, the offering of such cues by a person on video, in the absence of reciprocal interaction, usually was not sufficient for them to learn words. However, there is a second way that a video can direct attention to a labeled object: by presenting a *close-up* of the target (thus eliminating other potential referents) accompanied by a voiceover. Voices draw children's attention to a TV screen; preschoolers playing with toys in the presence of a television look at the screen at "noisy" moments such as when a woman speaks (Alwitt, Anderson, Lorch, & Levin, 1980). Research on word learning takes advantage of this fact: in the Intermodal Preferential Looking Paradigm (e.g., Golinkoff, Hirsh-Pasek, & Cauley, 1987), an object/action on video is labeled in a voiceover. After repeated pairings of the word and video, children are presented with the label accompanied by the old video and a new one showing a different object or action. Very young children tend to look to the matching video (e.g., Scofield, Williams, & Behrend, 2007; Golinkoff et al., 1987; Werker, Cohen, Lloyd, Casasola, & Stager, 1998), showing that they can learn associations between words and what they see on screen. Werker and colleagues (also Roseberry et al., in press) express the need for caution in describing such matching as

"word learning", reserving the term for situations in which children respond appropriately to a request (e.g., hand a questioner the object labeled with the novel word) or extend the label to other instances of the same category. At minimum, word learning from video would seem to require extending the label to the real object depicted on the screen (see Allen & Scofield, 2008).

Some children's television programs (e.g., *Blue's Clues, Dora the Explorer*) attempt to provide social cues, even a kind of contingency, by having the character on the screen ask questions and leave pauses for the child to answer. Research to date (involving preschoolers older than age 3) indicates that even though a person on television is not actually responsive, repeated viewing leads to more interaction on the part of the child, as well as increased comprehension of program content (Anderson, Bryant, Wilder, Santomero, Williams, & Crawley, 2000; Crawley, Anderson, Wilder, Williams & Santomero, 1999). By the age of 3, of course, children have been shown to learn words and other information from video (e.g., Rice, Huston, Truglio, & Wright, 1990; Rice & Woodsmall, 1988), possibly because they have begun to see videos (as well as pictures and scale models) as representations that can convey information (DeLoache, 2002; Troseth & DeLoache, 1998). We are currently conducting research to determine whether infants and toddlers also treat watching so-called "interactive" video as a learning situation.

By 30 months of age, toddlers have established scripts for the ways adults typically interact with them and teach them new information. They therefore have expectations about what learning situations are like, and may be more likely to learn in contexts that match their expectations. While scripts for pedagogical interactions may

vary across cultural contexts (Rogoff, 2003), the children we studied are likely used to teaching contexts that involve reciprocal (contingent) interactions with eye contact. The current research indicates that toddlers can learn words from observing these types of social interactions between others, both live and on video. That is, when they see knowledge being transferred between two people, whether in person or in a social interaction on a TV screen, they learn from the exchange, perhaps by imagining themselves as part of the interaction (Herold & Akhtar, 2008; Moore, 2007). Currently, it is not possible in commercial television to actually include the child in a reciprocal interaction with a teacher on the screen. Given the apparent advantage for very young children of learning from reciprocal interactions, television likely will remain at best an adjunct to live pedagogical interactions with involved adults.

REFERENCES

Akhtar, N. (2005). Is joint attention necessary for early language learning? In B.D. Homer & C.S. Tamis-LeMonda (Eds.), *The development of social cognition and communication.* (pp. 165-179). Mahwah, NJ: Erlbaum.

Akhtar, N., Jipson, J., & Callanan, M. (2001). Learning words through overhearing. *Child Development, 72,* 416–430.

Akhtar, N. & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology, 14,* 79-93.

Allen, R., & Scofield, J. (2008). Word learning from videos: Evidence from 2-year-olds. Manuscript submitted for publication.

Alwitt, L. F., Anderson, D. R., Lorch, E. P., & Levin, S. R. (1980). Preschool children's visual attention to attributes of television. *Human Communication Research, 7,* 52-67.

Anderson, D. R., Bryant, J., Wilder, A., Santomero, A., Williams, M., & Crawley, A. M. (2000). Researching *Blue's Clues*: Viewing behavior and impact. *Media Psychology, 2,* 179-194.

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist, 48*, 505 – 522.

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

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

Baldwin, D.A., & Moses, L.J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development, 10,* 309-329.

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior and Development, 19,* 159–170.

Barr, R. & Hayne, H. (1999). Developmental changes in imitation from television in infancy. *Child Development, 70,* 1067-1081.

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8,* 492-499.

Bigelow, A.E., MacLean, B.K., & MacDonald, D. (1996). Infants' response to live and replay interactions with self and mother. *Merrill Palmer Quarterly, 42*, 596-611.

Carpenter, M., Tomasello, M., & Striano, T. (2005). Role reversal imitation and language in typically developing infants and children with autism. *Infancy, 8,* 253-278.

Cooper, R. P., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61,* 1584-1595.

Crawley, A.M., Anderson, D.R., Wilder, A., Williams, M., & Santomero, A. (1999). Effects of repeated exposures to a single episode of the television program *Blue's Clues* on the viewing behaviors and comprehension of preschool children. *Journal of Educational Psychology, 91*, 630-637.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case of pedagogy. In M.H. Johnson & Y.M. Munakata (Eds.) *Processes of change in brain and cognitive development. Attention and Performance,* XXI (pp. 249-274). Oxford: Oxford University Press.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13,* 148-153.

DeLoache, J. S. (2002). The symbol-mindedness of young children. In W. Hartup & R. A. Weinberg (Eds.), *Child psychology in retrospect and prospect: In celebration of the 75th anniversary of the Institute of Child Development* (Vol. 32, pp. 73-101). Mahwah, NJ: Lawrence Erlbaum Associates.

Deocampo, J.A. & Hudson, J.A. (2005). When seeing is not believing: Two-year-olds' use of video representations to find a hidden toy. *Journal of Cognition and Development, 6,* 229-260.

Fenson, L., Pethick, S., Renda, C., Cox, J.L., Dale, P.S., & Reznick, J.S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21,* 95-116.

Floor, P., & Akhtar, N. (2006). Can 18-month-olds learn words by listening in on conversations? *Infancy, 9,* 327-339.

Golinkoff, R.M., Hirsh-Pasek, K., & Cauley, K.M. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*, 23-45.

Hains, S.M., & Muir, D.W. (1996). Effects of stimulus contingency in infant-adult interactions. *Infant Behavior and Development, 19,* 49-61.

Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. *Developmental Science, 6,* 254-261.

Herold, K.H., & Akhtar, N. (2008). Imitative learning from a third-party interaction: Relations with self-recognition and perspective taking. *Journal of Experimental Child Psychology, 101,* 114-123.

Király, I., Csibra, G., & Gergely, G. (2004). The role of communicative-referential cues in observational learning during the second year. Poster presented at the 14th biennial International Conference on Infant Studies, Chicago.

Koenig, M.A., & Echols, C.H. (2003). Infants' understanding of false labeling events: the referential roles of words and the speakers who use them. *Cognition, 87,* 179-208.

Krcmar, M., Grela, B.G., & Lin, Y.-J. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology, 10,* 41-63.

Kuhl, P.K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science, 10,* 110-120.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America (PNAS), 100,* 9096 – 9101.

Meltzoff, A. N. (1985). Immediate and deferred imitation in fourteen- and twenty-four-month-old infants. *Child Development, 56,* 62–72.

Meltzoff, A.N., & Moore, M.K. (2002). Imitation, memory and the representation of persons. *Infant Behavior and Development, 25,* 39-61.

Moore, C. (2007). Understanding self and others in the second year. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional development in the toddler years: Transitions and transformations* (pp. 43-65). New York, NY: The Guilford Press.

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology, 42,* 555-565.

Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science, 11,* 722-731.

Rice, M. L., Huston, A. C., Truglio, R., & Wright, J. (1990). Words from "Sesame Street": Learning vocabulary while viewing. *Developmental Psychology, 26,* 421-428.

Rice, M. L. & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development, 59,* 420-429.

Rideout, V. J., & Hamel, E. (2006). *The media family: Electronic media in the lives of infants, toddlers, preschoolers and their parents.* Menlo Park, CA: Kaiser Family Foundation.

Rogoff, B. (2003). *The cultural nature of human development.* Oxford: Oxford University Press.

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., Golinkoff, R.M. (in press). Live action: Can toddlers learn verbs from video? *Child Development.*

Ross, H.S., & Lollis, S.P. (1987). Communication within infant social games. *Developmental Psychology, 23,* 241-248.

Schmitt, K. L., & Anderson, D. R. (2002). Television and reality: Toddlers' use of visual information from video to guide behavior. *Media Psychology, 4,* 51 – 76.

Scofield, J., Williams, A., & Behrend, D.A. (2007). Word learning in the absence of a speaker. *First Language, 27,* 297-311.

Strouse, G.A., & Troseth, G. L. (2008). "Don't try this at home": Toddlers' imitation of new skills from people on video. *Journal of Experimental Child Psychology, 101,* 262-280.

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

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28,* 675-735.

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

Troseth, G.L., & DeLoache, J.S. (1998). The medium can obscure the message: Young children's understanding of video. *Child Development, 69,* 950 – 965.

Troseth, G.L., Saylor, M.M., & Archer, A.H. (2006). Young children's use of video as a source of socially relevant information. *Child Development, 77,* 786-799.

Troseth, G.L., Saylor, M.M., & Strouse, G.L. (2009). Do toddlers learn words from video? Unpublished manuscript.

Uzgiris, I.C. (1981). Two functions of imitation during infancy. *International Journal of Behavioral Development, 4,* 1-12.

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

Woodward, A.L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science, 6,* 297-311.