

This article was downloaded by: [Booth, Amy E.]

On: 6 August 2008

Access details: Access Details: [subscription number 794485673]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Language Learning and Development

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t775653671>

### Socio-Pragmatics and Attention: Contributions to Gesturally Guided Word Learning in Toddlers

Amy E. Booth <sup>a</sup>; Karla K. McGregor <sup>b</sup>; Katharina J. Rohlfing <sup>c</sup>

<sup>a</sup> Northwestern University, <sup>b</sup> University of Iowa, <sup>c</sup> Bielefeld University,

Online Publication Date: 01 July 2008

**To cite this Article** Booth, Amy E., McGregor, Karla K. and Rohlfing, Katharina J.(2008)'Socio-Pragmatics and Attention: Contributions to Gesturally Guided Word Learning in Toddlers',*Language Learning and Development*,4:3,179 — 202

**To link to this Article:** DOI: 10.1080/15475440802143091

**URL:** <http://dx.doi.org/10.1080/15475440802143091>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# Socio-Pragmatics and Attention: Contributions to Gesturally Guided Word Learning in Toddlers

Amy E. Booth  
*Northwestern University*

Karla K. McGregor  
*University of Iowa*

Katharina J. Rohlfing  
*Bielefeld University*

It is clear that gestural cues facilitate early word learning. In hopes of illuminating the relative contributions of attentional and socio-pragmatic factors to the mechanisms by which these cues exert their influence, we taught toddlers novel words with the support of a hierarchy of gestural cues. Twenty-eight- to 31-month-olds heard one of two possible referents labeled with a novel word, while the experimenter gazed at or gazed at and pointed to, touched, or manipulated the target. Learning improved with greater redundancy among cues, with the largest improvement evident when pointing was added to gazing. Looking times revealed that attentional factors accounted for only a small fraction of the variance in performance. Indeed, a significant increase in attention driven by manipulation of the target failed to improve learning. The results therefore suggest a strong role for socio-pragmatic factors in supporting the facilitative effect of gestural cues on word learning.

## INTRODUCTION

Although it is now well documented that gestural cues facilitate early word learning (Hollich et al., 2000), the relative contributions of different types of cues, and the mechanisms supporting their influence remain underspecified. What we do know is that as early as their first birthday infants appreciate pointing as an intentional act directed at something in the physical world (Woodward & Guajardo, 2002). Not only do 12-month-olds use pointing to indicate their own interests (Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004; Liszkowski, Carpenter, Striano, & Tomasello, 2006), but they are able to capitalize on others' points to localize specific objects in space (Morissette, Ricard, & Gouin Decarie, 1995; Thoermer & Sodian, 2001; Tomasello, Carpenter, & Liszkowski, 2007).<sup>1</sup> Of greatest relevance to the current work is Woodward's (2004) demonstration that pointing (in concert with eye gaze) specifically facilitates word learning in infants by 13 months. After one experimenter drew an infant's attention to a novel object, a second experimenter labeled it either while looking at and pointing to it, or while looking at a TV monitor. Infants were more likely to map a word onto its intended referent in the former condition.

By their second birthday infants respond to eye-gaze alone in a similar manner (e.g., Brand, 2000; Brooks & Meltzoff, 2005; Butler, Caron, & Brooks, 2000; Caron, Kiel, Dayton, & Butler, 2002; Corkum & Moore, 1998; Tomasello, 2001; Woodward, 2003). In a particularly striking demonstration, Baldwin (1993) found that 18-, but not 14-, month-old infants were sensitive to eye-gaze in a novel word learning task, even when it required them to disengage from their own current focus of attention. As one would expect, when the experimenter looked at and labeled objects to which infants were already attending, infants successfully mapped the words onto that focus of joint attention. However, if while the infant focused on one object the experimenter looked at and labeled a different object (that was hidden in a bucket), infants no longer consistently mapped words to the objects of their own attention. Rather, many of them mapped the words to the objects that were hidden from view during labeling.

Although the existing literature has made important contributions to our understanding of the role of gestural cues in supporting early word learning, it has so far been largely limited to the consideration of two specific gestures: gazing and pointing. What is lacking is a link to recent literature documenting the relevance of other nonverbal referential behavior (Brand, Baldwin, & Ashburn, 2002; Gogate, Bahrick, & Watson, 2000), as well as a direct comparison of

---

<sup>1</sup>The precise age at which this skill is evident actually varies with the angular displacement of the target from midline, its distance from the infant, and the hemisphere (left vs. right) in which it is presented (Butterworth & Itakura, 2000; Carpenter, Nagell, & Tomasello, 1998).

potential contributors to the word learning process. Addressing these weaknesses should lead to new insights at both descriptive and explanatory levels of analysis. In the current research, we therefore focus on three interrelated issues in need of greater specification in the literature.

First, we consider a range of gestural cues that children might utilize in learning new words. Although to date the focus of empirical inquiry has overwhelmingly been on looking and pointing, recent evidence suggests that other types of gestural cues are also available to children in the natural course of word learning. Gogate, Bahrick and Watson (2000) demonstrated that mothers label objects in synchrony with object motion (and sometimes touch) more often for words that they are explicitly trying to teach than for those they are not (see also Brand et al., 2002; Zukow-Goldring, 2006). Given that infants appreciate that touching and grasping are relevant to reference within their first year (Woodward, 1998; Woodward & Guajardo, 2002), it seems likely that early word learning would benefit from these pedagogical accommodations. However, there has been little experimental scrutiny of such an effect. In the only relevant study of which we are aware, Hennon et al. (2000) demonstrated that 12-month-olds could not learn a new word repeated five times on the basis of a touch (in concert with eye-gaze), but could learn a word presented 10 times on the basis of manipulation (in concert with eye-gaze).

Second, we consider the relative utility of different types of gestural cues for supporting early word learning. To date, direct comparisons between gestural cues under equivalent conditions have been rare. Again, Hennon et al. (2000) provide the most relevant data of which we are aware, and even here comparable conditions are tested against chance performance rather than against each other. In that work, 12-month-olds failed to learn new words (presented five times) regardless of whether the experimenter only gazed at, gazed at while touching, or gazed at while manipulating the target object. However, when words were presented 10 times, infants learned with the support of manipulation, whereas they still failed with gaze alone. This intriguing work begs numerous questions. Would differences in the effectiveness of more subtle shifts in gestural support (e.g., between gazing alone and gazing in conjunction with pointing) be evident under more supportive learning conditions? Would older children derive more benefit from these cues?

Third, we consider the mechanisms by which gestural cues support word learning. Attentional learning accounts suggest that gestural cues are helpful to the extent that they increase the salience of the intended referent, thereby supporting simple associations between simultaneously heard words and seen referents (Horst & Samuelson, *in press*; Samuelson & Smith, 1998; L. B. Smith, 1995). This perspective is supported by evidence demonstrating that the perceptual salience of potential referents can over-ride gestural cues within the first year of life and continues to influence learning well into the third year (Brand, 2000;

Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006; Samuelson & Smith, 1998; Smith, Jones, & Landau, 1996). An alternative possibility is that children have a rather sophisticated understanding of the referential intent of a variety of socio-pragmatic cues from a very early age, and that little more than this knowledge is necessary to learn new words in the natural course of experience (e.g., Tomasello, 1992, 2003; Tomasello et al., 2007). This perspective receives support from studies demonstrating the ability of infants as young as 18 months of age to use socio-pragmatic cues to infer the intended referents of novel words, even when those referents are not visible to them (e.g., Baldwin, 1993; Tomasello, Strosberg, & Akhtar, 1996).

The distinction between these perspectives of course is in reality not quite so clearcut. The vast majority of scientists working in this area suggest that both attentional and socio-pragmatic factors play important roles in word learning, albeit to different degrees and/or during different windows of development (e.g., Akhtar & Tomasello, 2000; Horst & Samuelson, in press; Moore, Angelopoulos, & Bennett, 1999; Tomasello, 2003; Woodward, 2004). In taking perhaps the most explicit stance on this issue, the emergentist coalition model weights perceptual-attentional factors heavily in the first year of life, but then describes a gradual shift towards the influence of socio-pragmatic factors over the following year (Hirsh-Pasek, Golinkoff, & Hollich, 2000). Although proponents of the emergentist coalition model argue that the dominance of socio-pragmatic factors is never complete, the role of attentional factors beyond this well-documented transition has not been well articulated, nor has it been empirically documented (Hollich et al., 2000).

One possibility is that *contiguity* between hearing a new word and attending to its intended referent continues to be helpful to young word learners. Evidence already demonstrates that this form of contiguity is not necessary for learning to occur in toddlers (e.g., Akhtar & Tomasello, 1996; Baldwin, 1993; Tomasello & Barton, 1994; Tomasello et al., 1996), and that it is not sufficient on its own to facilitate learning in infants under otherwise supportive conditions (Woodward, 2004). However, it remains possible that contiguity between the learner's auditory perception of the label and visual perception of the referent provides some modest level of support to the process.

Another possibility is that, contiguity aside, *total amount* of attention paid to the intended referent supports learning. After all, there is little question that time spent in focused attention translates into superior learning and retention of new information throughout development and across domains (e.g., Rovee-Collier & Hayne, 2000; Smith & Kosslyn, 2007). Memory for the properties of labeled objects would presumably enhance recognition on later encounters, and facilitate reactivation of associated information (i.e., labels). Evidence for this possibility is mixed. While Moore et al. (1999) found the number of times 24-month-old infants looked at the labeled target (in the context of two possible referents) to be

unrelated to word-learning performance, they did find a relationship between these measures in a subset of their 18-month-old sample. Moreover, they demonstrated that the older infants (but not the younger ones) could map novel words onto moving referents in the absence of any referential cues from the speaker, suggesting a continued role for target salience in learning. And although Hennon et al. (2000) conclude from their studies that “attention during labeling is not a good indicator of subsequent performance in test trials” (p. 83), they provide no direct test of this claim. In fact, a cursory examination of their summary data reveals that the only three (of eight) manipulations that led to successful learning in their 12-month-old subjects were precisely those that led infants to spend the most absolute amount of time attending to the target during training.

A final possibility is that basic perceptual-attentional factors come to play a truly negligible role in supporting word-learning sometime in the child’s third year of life. In their wake, the child’s attention to the word’s relationship to the intentions of the speaker and to the larger culture of communication might become paramount.

With these possibilities in mind, we taught 2½-year-old toddlers novel words in an unambiguous labeling context with the support of a hierarchy of gestural cues. Some children saw the experimenter merely gaze at one of two potential referent objects while providing a novel label (G condition). Others saw her gaze at and point to it (GP condition). Others saw her gaze at and extend her point so that she ultimately touched the target object (GT condition). Others saw her gaze at, touch, and manipulate the target object by pushing it across the table (GM condition). We explicitly chose this embedded hierarchy of referential cues (rather than pitting individual cues against each other) in order to best represent the types of labeling episodes that young children naturally experience. In a final baseline control condition, toddlers received no referential cues during labeling (BL condition). Immediately, and again several days after training, children were tested on their production, comprehension, and extension of the novel words. Attention at the time of labeling and throughout training was also assessed.

Based on evidence suggesting that gestural cues (including eye-gaze alone) are useful to much younger infants in learning new words (e.g., Baldwin, 1993; Woodward, 2004) we expected all levels of the gestural hierarchy to facilitate word learning. We considered the possibility that, given the relatively mature status of our participants, all of the conditions would in fact be equally effective in this regard. By 2½ years of age, children might readily recognize the familiar “naming context” (Campbell & Namy, 2003; Namy, 2001) and efficiently scan the speaker for even subtle cues (i.e., eye-gaze) to the location of the intended target. Their performance might therefore be at or near ceiling across all conditions. We suspected, however, that this would not be the case. Instead, we predicted that differences in the effectiveness of conditions would emerge and that

these differences would be revealing about the mechanisms responsible for the facilitative effect of gestural cues on early word learning.

For example, if attention to the intended referent (either contiguously with hearing the label or not) is an important contributor, then those conditions that best corral attention in this manner should also best facilitate word learning. This will be most successfully achieved by gestures that are readily detected and that efficiently isolate the spatial location of the target. This becomes more characteristic of the gestural cues provided as we advance up the proposed hierarchy and as more body parts, more extensive motions, and more complete specification of the linear trajectory from speaker to referent are invoked. Although the GT and GM conditions differ minimally on these factors, they differ in another way that is perhaps even more relevant. By including physical manipulation of the target, the GM condition should engage the ubiquitous attentional pull of motion (Dannemiller & Freedland, 1993; Kellman, 1993) thereby directing the infant's focus to the critical spatial location (Moore et al., 1999; (also see Gogate & Bahrick, 1998; Gogate & Bahrick, 2001; Hennon, 2000; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). If attention to the target does play a central role in facilitating word learning, one might expect to see the greatest difference across consecutive levels of the gestural hierarchy in comparison of the GT and GM conditions. Further, attention to the target during labeling and/or throughout training should be related to word-learning performance.

If, on the other hand, attention to socio-pragmatic context is an important contributor, then gestural cues that most uniquely and robustly tap into the child's concept of communicative intent should best facilitate learning. Because pointing occurs almost exclusively within referential contexts, while gazing occurs in much more variable circumstances, the former should be more effective than the latter in this regard. Providing additional cues (i.e., touching and manipulating) might be helpful to the extent that they reinforce the advantageous interpretation of communicative intent initiated by the point, but the most dramatic difference between consecutive levels of the hierarchy should emerge between the G and GP conditions. In contrast to the predictions set forth on the basis of attention to the target, no difference at all should be evident between the GT and GM conditions, which might be construed as equally informative about the referential intent of the speaker.

In sum, the existing literature suggests that eye-gaze presented either alone, or in combination with other gestural cues, should facilitate word learning by approximately 2 years of age. However, we currently know little about the relative utility of gazing, and gazing in conjunction with pointing, touching, or manipulating for young word learners. We know particularly little about the child's utilization of touching and manipulation as cues to reference (Brand et al., 2002; Gogate et al., 2000; Zukow-Goldring, 2006). In addition, despite widespread agreement that attentional factors continue to contribute to word learning

beyond the first year of life, we have little understanding of its lasting influence. In an attempt to further build our understanding in these areas, we ask whether a hierarchy of gestural cues will facilitate acquisition in 2 ½-year-old children. We expect that all levels of the hierarchy will facilitate learning, but to varying degrees. We consider the specific relation of gestural cues to “referential intent” as well as their ability to direct attention to the target as potential factors moderating their influence.

## METHOD

### Participants

Participants were 80 28- to 31-month-old toddlers ( $M = 29.31$ ,  $sd = .89$ ) being raised with English as their primary language (i.e., were exposed to less than 50% of an alternative language). An additional eight toddlers participated, but were subsequently excluded due to experimenter error ( $n = 3$ ), equipment failure ( $n = 2$ ) or failure to complete more than half of the procedure ( $n = 3$ ). There were a total of 38 females in the final sample (six in the baseline, nine in the gaze, eight in the point, nine in the touch, and seven in the manipulate condition). The sample included primarily Caucasian toddlers from the northern suburbs of Chicago.

### Materials

Stimuli for the teaching episodes were three pairs of novel objects. Each pair included a target object (e.g., a carpenter’s level) and a foil (e.g., a soap basket). These same stimulus pairs were also presented in the production test. For the comprehension test these stimulus pairs were embedded in a larger set of novel objects that included an additional exemplar drawn from the same category as the target (e.g., a different carpenter’s level), an additional exemplar drawn for the same category as the foil (e.g., a different soap basket), and a fifth object drawn from yet a third novel category (e.g., tongue scraper). The purpose of the additional target exemplar was to determine whether the child had not only mapped but also extended the new label to other relevant category members. The purpose of the additional foil exemplar was to avoid cueing the child’s extension response by presenting only one “large” category in the array. Finally, the purpose of the third category exemplar was to decrease chance responding to 20%. Furthermore, responses involving this previously unseen third category would indicate a novelty strategy and would counter-indicate task comprehension. The three complete sets of experimental stimuli are presented in Fig. 1.

Two sets of more familiar stimuli were also used for training on the comprehension test. One set included a shoe, a bowl and three different spoons. The



FIGURE 1 The 3 sets of stimuli presented to all participants. The target for each set is pictured in the left-most column.

other included two different books, a dog, a fish, and a novel oval-shaped object. The purpose of the multiple spoons and books was to teach the children that more than one “correct” response was acceptable. *Three* spoons, but only *two* books, were included to give the child permission to give a variable number of “correct” responses. Finally, a novel object was included in one set so that the child was not thrown by the novelty of objects in the actual test arrays.

## Design

Toddlers were randomly assigned, 16 per group, to a gaze (G), gaze+point (GP), gaze+touch (GT), or gaze+manipulate (GM) condition, by which the referential cues available to toddlers for determining mappings between objects and labels varied. An additional 16 infants were run in a baseline (BL) condition by the same experimenters after all other conditions were completed. This condition was added in order to discount the possibility that above chance performance in any condition resulted from baseline preferences for the exemplars of the target category. All toddlers saw precisely the same stimuli throughout the procedure and received the same number of exposures to novel object labels.

## Procedure

An overview of the procedure appears in the Appendix.

After a brief warm-up period of free play with the experimenter, toddlers were seated in a high-chair pushed up to a table. The experimenter sat across from the

toddler. The parent sat behind the child and was asked not to interact during the procedure but instead to complete the MacArthur-Bates Communicative Development Inventory (Fenson et al., 1993).

*Word teaching episode 1.* The experimenter placed target-foil pair one on the table in front, but out of reach of the toddler. The objects were placed one slightly to the right and one slightly to the left of the experimenter, approximately 40 cm apart, and equally distant from the toddler. The experimenter then labeled the target object three times (e.g., “Look, it’s a koob! It’s called a koob! Wow, it’s a koob!”). While labeling, the experimenter performed the following condition-specific action. In the G condition she turned her head and gazed at the target. In the GP condition she turned her head, gazed at the target, and pointed to it with her right index finger, leaving approximately seven inches between her finger and the target. In the GT condition she turned her head, gazed, and pointed to the target as in the GP condition, but extended her arm until her finger was touching the target. In the GM condition she turned her head to gaze and touch the target as in the GT condition, but then used her index finger to push the object slowly across the table. Half of the children in this condition saw the target object moving approximately 14 cm away from the foil; the other half saw the target moved the same distance towards the foil. The purpose of this manipulation was to ensure that any effects in the GM condition were due to movement of the target as opposed to changes in the proximity between the target and the foil. In the BL condition the experimenter simply looked at a point on the table midway between the target and foil while labeling.

The experimenter then gave the object pair to the child to play with for 10 seconds. After retrieving the objects, the experimenter reversed their original placement on the table and repeated the labeling episode just described.

The procedure described above was then repeated for target-foil pairs two and three. The specific object pairs that were presented first, second, and third varied randomly across children.

*Choice task training.* This training was designed to familiarize children with the comprehension test procedure. The experimenter handed one set of training objects to the child and said, “Look at these!” She allowed the child to play with the objects for 30 seconds. She then said: “Ok, I’m going to scoop everything up so we can play a game. Now I’m going to line everything up on the tray.” After lining up the objects in random order, the experimenter introduced a dog puppet and said, “Look who I have . . . it’s Doggy! Listen carefully. Doggy says that she only likes spoons/books. Are there any spoons/books here for Doggy? If you see a spoon/book, give it to Doggy.” The experimenter then pushed the tray within the child’s reach. Once the child handed one item to the experimenter, she asked: “Are there any more spoons/books for Doggy?” The experimenter repeated this request until all of the correct items were handed to

her. If the child made a mistake, she said: "Oops, that's not a spoon/book, so we will leave it on the tray. Are there any more *spoons/books* for Doggy?" This procedure was repeated with the second set of training objects.

*Word teaching episode 2.* The procedures of Word Teaching Episode 1 were repeated, except only one of the three target-foil pairs was introduced before proceeding with testing as described below. The teaching and testing episodes were then repeated for the second and third target-foil pairs, with order of the specific pairs varying randomly across children.

In sum, prior to testing children had four opportunities (two per teaching episode) to view each pair of objects along with condition-specific gesture cues that indicated which member of the pair was the intended referent. Given that the experimenter named each referent three times during each of these opportunities, the children received a total of 12 exposures to each label.

*Production testing.* The experimenter placed the target and foil on the table and introduced a new stuffed animal. She said: "Oh look, it's Bunny. She says these things look really neat. She wants to know what they are called. Can you tell her?" Then she held up either the target or foil that was presented during training and asked, "Do you know what this is called? . . . What is it?" Regardless of the child's response, she said "ok." The experimenter then repeated her query with the remaining object. The child was not expected to know the name of the foil because it was not labeled during training. This query was only included so as not to differentially highlight the target prior to testing.

*Comprehension testing.* The child was first allowed to play freely with all five test objects for 30 seconds. The experimenter then lined up the objects in random order on a tray out of the child's reach and said, for example, "Bunny says she only likes koobs. Are there any koobs here for her? If you find a koob, give it to Bunny." Then she pushed the tray within the child's reach and waited for a response. If the child offered an object, she said, "Thank you and asked, "Are there any more koobs for Bunny?" This was repeated up to two additional times. The experimenter stopped if at any point the child responded, "No," to her query. To ensure that all participants heard the target words the same number of times, if a child responded without extra prompts, the experimenter repeated the word in neutral contexts (e.g., Bunny really likes koobs) to compensate.

Word teaching episode 2, along with the subsequent production and comprehension testing sessions, was repeated for each of the two remaining stimulus sets.

*Follow-up testing.* All of the toddlers were scheduled for a follow-up session three to five days later, at which they were tested for their comprehension and production of the target words as in the initial testing phase.

*Coding.* Digitized videos of the experimental sessions were coded by trained observers using Noldus Observer software. Coders were blind to the experimental hypotheses and condition assignment of individual participants. For each videotape a primary coder transcribed all target word productions and recorded the order in which test objects were selected during comprehension testing. For all tapes on which the toddler's eyes were fully visible (all but two), they also coded the time spent looking at the target (in contrast to the experimenter and foil) during the observation periods of the word teaching episodes. A second coder scored 25% of the videotapes equally sampled from across conditions. Average percent agreement among coders was 100% for the sequence of test object choices and 93% for patterns of looking ( $\kappa = .85$ ).

## RESULTS

We assigned each child a production score based on the following criteria: one point for each perfect production of the word, plus half a point for each partial production that included at least two of the three correct phonemes (e.g., *teep* instead of *teeg*). Half a point was deducted if the child also used the word to refer to the foil, or if they used one of the trained words, but applied it to the target object from the wrong set. As can be seen in Table 1, on average children produced less than one word per testing session. Nevertheless, we compared each experimental condition to baseline. These analyses revealed no significant differences. We also submitted the data to a repeated measures ANOVA that included session (First vs. Follow-up) as the within subjects factor and condition (G vs. GP vs. GT vs. GM) as the between subjects factor (see Table 1). A main effect of session revealed superior performance on the first, rather than the follow-up test,  $F(1, 60) = 14.67, p < .001$ . However, no effect of condition was evident. Univariate tests of performance at each testing session also revealed no effect of condition. We, therefore, conducted no further analyses on the production data.<sup>2</sup>

We calculated both a liberal and a stringent measure of comprehension. Our liberal measure reflected successful mapping of the novel words to either the

---

<sup>2</sup>It is unclear why production scores were so low in this task. It is possible that the toddlers' willingness to speak up might have simply been depressed by the novelty of the experimenter and setting. Alternatively, toddlers might have been unable to produce the words even under the best of testing circumstances. Evidence suggests that a deeper level of learning is required to support production as compared to comprehension (McGregor, Newman, Reilly, & Capone, 2002). Toddlers might have required a more extensive training procedure than that provided here (including, perhaps, opportunities to practice imitating the productions modeled by the experimenter), in order to produce the words independently at test.

TABLE 1  
 Children's Mean Comprehension and Production Scores (with standard errors)  
 in the Five Conditions at Initial Test and Follow-up

Condition	First session		Follow-up session			
	Comprehension		Production	Comprehension		Production
	Liberal	Stringent		Liberal	Stringent	
BL	.39 (.07)	.18 (.05)	.69 (.25)	.39 (.08)	.14 (.05)	.63 (.19)
G	.5 (.08)	.42 (.08)	.72 (.21)	.52 (.09)	.38 (.09)	.44 (.16)
GP	.69 (.08)	.56 (.10)	1.09 (.21)	.69 (.09)	.60 (.08)	.69 (.19)
GT	.76 (.08)	.70 (.08)	1.16 (.23)	.80 (.06)	.74 (.09)	.63 (.15)
GM	.77 (.07)	.69 (.08)	.84 (.24)	.74 (.09)	.63 (.10)	.63 (.20)

Note. Comprehension scores reflect proportions of correct choices on the child's first response (i.e., liberal; chance = .40) and both their first and second responses (i.e., stringent; chance = .10). Production scores reflect adjusted number of correct responses (range = 0 to 3).

target or to the new object drawn from that same category. For each child we calculated the proportion of such responses across the three stimulus sets. Our more stringent measure was calculated as the proportion of trials on which children selected the two target category exemplars as their first and second choice. Unlike the liberal measure, this stringent measure required that infants both map and extend the novel word to its appropriate referents. Both measures were calculated independently for the initial and follow-up testing sessions.

First, we asked whether all levels of the hierarchy were effective in facilitating word learning above the baseline condition in which performance was statistically equivalent to chance (i.e., .40 for the liberal, and 10 for the stringent measure).<sup>3</sup> As can be seen in Table 1, our liberal measure revealed superior comprehension in the GP, GT, and GM conditions than in the BL condition at both testing sessions ( $t(30)$  ranged from 2.55 to 4.16, all  $ps < .01$ ). Performance in the G condition did not differ from baseline in this analysis for either the first

<sup>3</sup>There is some possibility that the pragmatically bizarre behavior of the experimenter during the labeling episodes of the baseline condition (i.e., looking down at the table) either distracted the toddlers or led them to believe that neither the target nor the foil could be the referent of the novel word. However, a careful examination of the data collected in this condition suggests that this was not the case. First, toddlers spent 47% of the training episodes looking at the experimenter, a proportion that is very much in line with those observed in the other conditions. If toddlers were very surprised or confused by the experimenter's behavior, they should have spent more time looking at the experimenter in an attempt to figure out what she was doing. Second, toddlers chose all of the objects with approximately equal frequency at test (target=12, target extension=10, foil=10, foil extension=10, novel=8). If toddlers were rejecting the target and foil as potential referents of the novel word, they should have selected these objects less frequently than the others.

or second session,  $t(30) = 1.09$ , *ns*, and  $t(30) = 1.15$ , *ns*, respectively. However, examination of our more stringent measure, suggests a stronger effect of gaze on word learning,  $t(30) = 2.60$ ,  $p < .025$  and  $t(30) = 2.41$ ,  $p < .025$ , for the first and second session respectively. Comprehension performance according to the stringent measure was also higher than baseline in the GP, GT, and GM conditions ( $t(30)$  ranged from 3.60 to 5.89, all  $ps < .01$ ). Importantly, these differences cannot be attributed to preferences that might have developed for the target during training as a result of exposure to the directive gestures alone. During the 30 seconds of free-play with all of the test objects that immediately preceded forced-choice testing, attention to the target object alone, and to both members of the target category combined, was equivalent to baseline in each experimental condition,  $t(30)$  ranged from .01 to 1.35, all  $ps$  *ns*.

Next we asked whether the conditions differentially affected performance. We submitted each of our measures of comprehension (i.e., liberal and stringent) to separate repeated measures ANOVAs that included session (First vs. Follow-up) as a within subjects factor and experimental condition (G vs. GP vs. GT vs. GM) as a between subjects factor. No effect of session was evident for either measure, suggesting that children were able to retain whatever information they acquired over the course of several days. As predicted, a main effect of condition was evident for both the liberal,  $F(3, 60) = 3.47$ ,  $p < .02$ ,  $R^2 = .28$ , and the stringent,  $F(3, 60) = 3.44$ ,  $p < .02$ ,  $R^2 = .36$ , measures, indicating that some combinations of gestures were more facilitative than others.<sup>4</sup> Bonferroni post-hoc comparisons revealed differences between the G and GT and between the G and GM conditions only,  $ps < .01$ . However, in order to specifically test our alternative hypotheses, we also conducted a series of planned comparisons across consecutive levels of the gestural hierarchy. One-tailed t-tests were licensed by clear predictions for superior performance emerging as additional cues are added. Decrements in performance would not be predicted either by extreme socio-pragmatic or attention based accounts of word learning. When collapsed across session, the largest difference between consecutive levels of the hierarchy, and indeed the only difference that was significant, emerged in comparison of the G to the GP condition,  $t(30) = 1.86$ ,  $p < .05$ ,  $g = .65$  for the liberal measure and  $t(30) = 1.71$ ,  $p < .05$ ,  $g = .57$  for the stringent measure.<sup>5</sup> Effect sizes derived from

<sup>4</sup>These main effects of condition were maintained in parallel analyses including percentile rank on the MBCDI as a covariate. Scores on the MBCDI did correlate with both measures of comprehension, but did not account for a significant amount of variance in these analyses. MBCDI scores also did not differ across conditions, (BL:  $M = 63$ ,  $se = 8.9$ ; G:  $M = 65$ ,  $se = 6.8$ ; P:  $M = 61$ ,  $se = 7.3$ ; GT:  $M = 75$ ,  $se = 6.8$ ; GM:  $M = 64$ ,  $se = 6.3$ ) and were therefore not considered further in subsequent analyses.

<sup>5</sup>There was also no difference between the move-together and move-apart versions of the GM condition.

comparison of GP to GM were particularly small,  $g = .04$  for the liberal measure and  $g = -.20$  for the stringent measure. See Fig. 2.<sup>6</sup>

Next, we asked whether these differences in effectiveness were related to toddlers' simultaneous perception of the label and referent during training. For each toddler we calculated the proportion of labeling episodes (out of a maximum of 36) on which they looked at the target object while the experimenter produced its label (i.e., instances of cross-modal perceptual simultaneity). A significant bivariate correlation between condition (including all five levels) and cross-modal perceptual simultaneity was evident,  $r = .69$ ,  $p < .001$ . Toddlers in the BL condition looked at the target during labeling infrequently ( $M = .17$ ,  $se = .02$ ). All four combinations of gestural cues increased the likelihood of cross-modal perceptual simultaneity

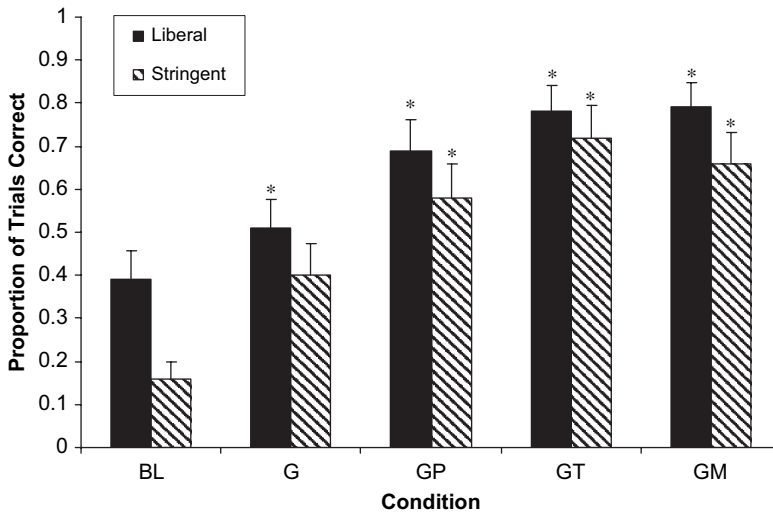


FIGURE 2 Mean liberal (chance = .40) and stringent (chance = .10) comprehension scores and standard errors for each condition collapsed across testing session. A \* indicates performance significantly above that observed in the baseline condition.

<sup>6</sup>These results were largely mirrored in an examination of individual patterns of performance. Overall, the distribution of toddlers revealing a perfect response pattern according to our liberal measure (i.e., choosing a target category exemplar first on all three tests) varied by condition at the first ( $X^2(3, N = 64) = 8.15$ ,  $p = .04$ ) testing session. In fact, there was but one perfect responder in the G condition, in contrast to five in the GP, seven in the GT and eight in the GM condition. A marginally significant effect of condition was also evident at the second ( $X^2(3, N = 64) = 7.62$ ,  $p = .06$ ) testing session with three perfect responders in the G, 6 in the GP, nine in the GT and 10 in the GM condition. However, our more stringent measure only revealed a significant effect at the follow-up testing session ( $X^2(3, N = 64) = 8.73$ ,  $p = .03$ ).

above this baseline,  $t(30)$ s ranged from 3.59 to 7.45, all  $ps < .001$ ,  $gs > 1.21$ . Comparisons of consecutive levels of the hierarchy revealed no differences between the G ( $M = .28$ ,  $se = .02$ ) and GP ( $M = .32$ ,  $se = .04$ ) conditions,  $g = .27$ . However, the GM condition ( $M = .59$ ,  $se = .05$ ) produced more cross-modal perceptual simultaneity than did the GT condition ( $M = .41$ ,  $se = .04$ ),  $t(30) = 2.70$ ,  $p = .01$ ,  $g = .92$ . A trend towards more cross-modal perceptual simultaneity in the GT than the GP condition also emerged,  $t(30) = 1.84$ ,  $p = .08$ ,  $g = .64$ . More distal levels of the hierarchy also differed from each other, all  $ps < .01$ ,  $gs > .99$ .

Next, we asked whether differences in the effectiveness of conditions were related to the total amount of attention that was directed to the target in each during training. We began by calculating the proportion of time during the observational portions of the word-teaching episodes (i.e., the times when toddlers were not able to touch the objects) that toddlers were looking at the target (in contrast to the experimenter or the foil). A significant bivariate correlation between condition (including all 5 levels) and proportion of time spent looking at the target was evident,  $r = .73$ ,  $p < .001$ . As can be seen in Fig. 3, toddlers in the BL condition spent most of their time looking at the experimenter ( $M = .47$ ,  $se = .06$ ), and equally small proportions of time looking at the target ( $M = .16$ ,  $se = .05$ ) and foil ( $M = .17$ ,  $se = .03$ ). All four combinations of gestural cues increased infants'

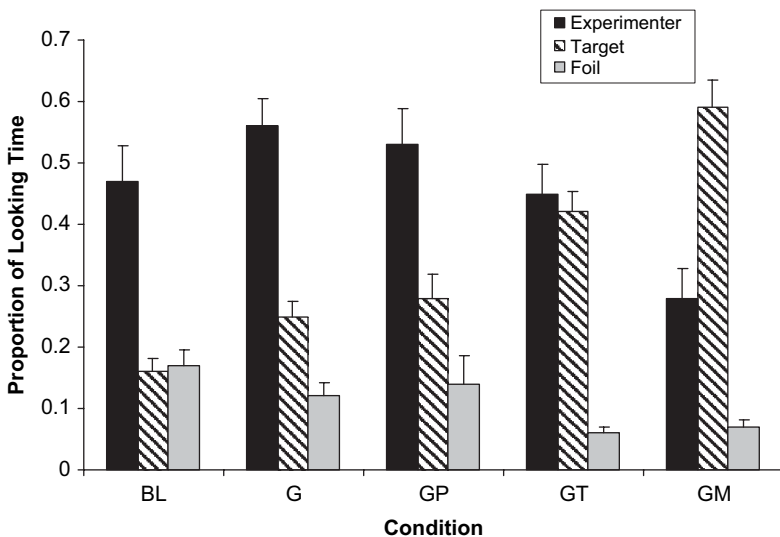


FIGURE 3 Mean proportions and standard errors of time toddlers spent looking at the experimenter, target and foil in each condition during the observation portions of the word-teaching episodes.

attention to the target above this baseline,  $t(30)$ s ranged from 2.55 to 8.24, all  $ps < .02$ . Comparisons of consecutive levels of the hierarchy revealed no differences between the G and GP conditions,  $g = .23$ . However, the GM condition ( $M = .59$ ,  $se = .05$ ) directed more attention to the target than the GT condition ( $M = .42$ ,  $se = .03$ ),  $t(30) = 3.08$ ,  $p < .01$ ,  $g = 1.06$ , which in turn directed more attention than the GP condition ( $M = .28$ ,  $se = .04$ ),  $t(30) = 2.91$ ,  $p < .01$ ,  $g = .97$ . More distal levels of the hierarchy also differed from each other, all  $ps < .001$ ,  $gs > 1.4$ .

Finally, we asked whether attention patterns during the word-teaching episodes were directly related to learning. When all conditions were included, cross-modal perceptual simultaneity correlated with both the liberal,  $r = .37$ ,  $p < .01$ , and stringent,  $r = .32$ ,  $p < .01$  measures of comprehension. Total looking to the target also correlated with both the liberal,  $r = .28$ ,  $p < .02$ , and stringent,  $r = .22$ ,  $p < .05$  measure. Notably, however, correlations between looking to the target and each measure of comprehension approached zero (both  $rs < .03$ , ns) when the effect of cross-modal perceptual simultaneity was partialled out of the analysis. In contrast, the correlations between cross-modal perceptual simultaneity and measures of comprehension remained at or near significance ( $r = .25$ ,  $p = .03$  for the liberal, and  $r = .22$ ,  $p = .057$  for the stringent, measure) when total looking to the target was partialled out of the analysis. Importantly, correlations between condition and measures of comprehension remained strong even when both measures of attention were simultaneously partialled out of the analyses,  $r = .41$ ,  $p < .001$  for the liberal, and  $r = .56$ ,  $p < .001$ , for the stringent measure.

## DISCUSSION

To review, toddlers were taught three novel count nouns with the support of one of four levels of a hierarchy of gestural cues; gaze alone, gaze+point, gaze+touch, or gaze+manipulate. Given the age of our participants, and the clear referential context of our naming episodes, we expected all forms of gestural support to facilitate word learning above a baseline condition. However, we did not expect all levels of the hierarchy to be equally effective. We formulated a set of contrasting hypotheses to specify how performance might vary across conditions if attention to (1) the target or (2) the broader socio-pragmatic context play a critical role in supporting word learning. We reasoned that if attention to the target (either contiguously with the label or not) was critical, toddlers' performance should improve as they are exposed to more elaborate gestures that more explicitly indicate the location of the intended referent (i.e., as they advance up the hierarchy of cues). The most dramatic improvement should therefore occur on the final step (from the GT to the GM condition) when the target is moved and the fundamental bias to attend to dynamic targets is engaged. In contrast to this set of predictions, if the socio-pragmatic relevance of the gestures play a critical

role, toddler's performance should be optimized when supported by combinations of gestural cues that include pointing (because of its unique relationship to referential intent). The greatest shift in performance across hierarchical levels would therefore be expected in comparison of the G and GP conditions. Although providing redundant cues like touching or manipulating might be expected to improve learning by reinforcing the interpretation of intentionality engendered by pointing, little or no difference in performance should emerge in comparison of the GT and GM conditions.

As predicted, cues comprising each level of the hierarchy facilitated the acquisition of novel words, as well as their retention over an extended delay.<sup>7</sup> These findings contribute to our understanding of gesture as a support for learning broadly defined (e.g., Church, Ayman-Nolley, & Mahootian, 2004; Valenzeno, Alibali, & Klatzky, 2003). More specifically, they converge well with evidence documenting the ability of infants and young children to map and retain novel words presented in the context of ostensive labeling (Waxman & Booth, 2000; Woodward, Markman, & Fitzsimmons, 1994), despite the fact that 2-year-olds have difficulty with retention in the absence of such supports (Horst & Samuelson, *in press*). At a still finer level of analysis, the current results build on research demonstrating that 13-month-olds are more likely to map a novel word onto its intended referent when the speaker points at it than when they do not (Woodward, 2004). In the current work, substantially older 2½-year-olds were able to use pointing to disambiguate an intended referent from two simultaneously presented alternatives. Similarly, our results extend conclusions regarding the facilitative effect of touching and manipulation on word learning in younger infants (e.g., Brand, 2000; Hollich et al., 2000) to toddlers. Touching and/or manipulating a referent object in concert with gaze resulted in better word learning than gazing alone. These findings are consistent with recent evidence suggesting that parents use both of these cues naturalistically in teaching their young children new words (e.g., Gogate & Bahrack, 1998; Gogate et al., 2000; Zukow-Goldring, 2006).

Results regarding the relative utility of different combinations of cues were more consistent with predictions derived from socio-pragmatic than attentional learning theory, emphasizing a heavy influence of attention to the broader communicative context over attention to the target object *per se*. Although performance generally improved as children advanced up the gestural hierarchy, the

---

<sup>7</sup>The careful reader will remember that performance on the stringent, but not the liberal measure, exceeded baseline in the gaze condition. Given that the former is logically entailed in the latter, and that previous research has shown that 24-month-olds can use gaze to map novel words to their referents (Brand, 2000), we are inclined to interpret the negative component of our findings regarding gaze as anomalous.

only significant difference across consecutive levels was evident in comparison of the G and GP conditions.

It is theoretically possible for this pattern of results to be accounted for by lower-level attentional factors. Indeed, the high correlations observed between condition and measures of attention to the target suggest that the effects of the former might well be mediated by the latter. It is easy to see how this might be so. As predicted, combinations of gestural cues that (1) involved more extensive motions, (2) provided a concrete indication of the linear path between speaker and referent, and (3) culminated in motion of the target did increase attention to the target during training. Despite this seemingly tight relationship, however, gesture cue condition explained much of the variance in performance, even after cross-modal perceptual simultaneity and total looking to the target were statistically accounted for. Moreover, this attention-based explanation is not consistent with the fact that the G and GP conditions, which differed in performance, did not also differ in the degree to which they directed attention to the target object during labeling or throughout training. Word-learning performance also did not improve across conditions where a difference in attention to the target was observed (i.e., GP vs. GT and GT vs. GM).

Nevertheless, the data suggest that attention to the target does play some role in facilitating acquisition. Both cross-modal perceptual simultaneity at the time of labeling and total time spent looking at the target correlated strongly with condition. Moreover, the conditions that differed most dramatically in performance (i.e., G vs. GT and G vs. GM) also differed most dramatically on both measures of attention to the target. Both measures of attention to the target also correlated weakly with both the liberal and stringent measures, thereby accounting for small, but significant amounts of the variance in comprehension.

Interestingly, correlations between cross-modal perceptual simultaneity and measures of comprehension remained significant after controlling for the influence of time attending to the target during training, but not vice versa. This suggests that the total amount of time that toddlers attend to the target during training only influences learning to the extent that it increases the likelihood that toddlers will be focused on the target at the time of labeling. The importance of simultaneity of experience broadly speaking resonates with associationist theories founded on Hebbian learning mechanisms (Elman, 2001; L. B. Smith, 1995). According to such accounts, the frequency with which two events co-occur (e.g., experience of label and visual referent) directly determines the likelihood that memory for one (e.g., the label) will be activated when the other is again encountered (e.g., the visual referent). However, the fact that total time attending to the target did not directly influence word learning is somewhat surprising in light of the ubiquitous role attributed to focused attention in learning (E. E. Smith & Kosslyn, 2007). And while our failure to find a strong relationship between total inspection time and learning is consistent with several existing studies in the

specific area of early word learning (e.g., Brand, 2000; Campbell & Namy, 2003; Woodward, 2004), it is worth noting that this factor might have its greatest impact on the *longevity* of memories and therefore might become more predictive of comprehension after lengthier testing delays. Given that performance was equally strong at the time of initial and delayed test in the current work, we clearly failed to provide a sensitive test of this possibility. Testing after a longer delay may be more revealing.

Given the relatively small contribution of attentional factors to explaining word learning outcome, what else might account for the differences observed across conditions? Although the current investigation cannot provide a clear answer to this question, it seems worthwhile to speculate on some possibilities. As previously suggested, one possibility is that the conditions vary in the degree to which they encourage the learner to interpret the speaker's behavior in terms of referential intent. Cues that are tightly associated with this type of communicative exchange in the child's experience (i.e., pointing) might provide them with the strongest justification and motivation for linking the word to its referent, and might most successfully promote their joint attention (in the rich sense of intersubjectivity) with the adult speaker (Baldwin, 1995; Tomasello, 1995; Tomasello et al., 2007). It is also possible that by heightening attention to the broader socio-pragmatic context, gestural acts that include pointing engender more elaborated, deeper, and/or robust encoding of associations between words and referents. It is well known that elaboration of representations leads to more robust memories, particularly when that elaboration has a meaningful or causal structure—such as that which might be supported by attention to the goals and interests of the speaker (see Bauer, 1992; Bradshaw & Anderson, 1982; Capone & McGregor, 2005; Cople & Coon, 1977; E. E. Smith & Kosslyn, 2007). Word learning might be expected to reap particular benefit from this type of elaboration given its tight relationship with conceptual knowledge from early in development (e.g. Booth & Waxman, 2002; Gelman & Markman, 1987; Graham, Kilbreath, & Welder, 2004).

## CONCLUSIONS

In conclusion, the current research demonstrates that a variety of gestural cues facilitate word learning in 2 ½-year-olds. In particular, it provides clear evidence of a positive influence of gazing, pointing, touching, and manipulating (all in concert with gaze) on the ability of toddlers to isolate the intended referents of novel words. The current research also highlights that all combinations of gestural cues are not equally effective in supporting this process. Conditions that provided a confluence of redundant cues were most effective, with the most dramatic gain

emerging in transition from gaze alone, to pointing in concert with gaze.

Although performance could be partially accounted for by attentional factors (e.g., simultaneous attention to the label and target), the influence of condition was far greater, thereby suggesting a strong role for socio-pragmatic factors (e.g., relevance to communicative intent) as well. As predicted by proponents of the Emergentist Coalition Model (Hirsh-Pasek et al., 2000), it appears that socio-pragmatic factors come to play a larger role than perceptual-attentional factors in word learning by the time children reach 2 ½ years of age (see Gogate, Walker-Andrews, & Bahrick, 2001 for additional evidence consistent with this developmental progression away from a reliance on low-level perceptual-attentional factors; Moore et al., 1999).

Although these are important new contributions, much work remains to be done in clarifying the relative contributions of different combinations of gestural cues to word learning both across development and learning contexts. It seems likely that their differential effectiveness will change as children develop skills and expectations that improve their detection of the cues (regarding eye gaze in particular), and as they learn their culturally prescribed importance (regarding pointing in particular). In addition, it seems likely that the relative usefulness of the cues will vary depending on the lexical skills of the learner, as well as environmental factors such as the distal proximity of target objects and the degree of clutter surrounding them. Therefore, we fully expect that applying a wider empirical lens will reveal an intricate developmental story. As it unfolds, this story should provide further clues to the underlying mechanisms responsible for the facilitative effect of providing gestural cues on early word learning.

## ACKNOWLEDGMENTS

This research was supported in part by an internal grant from Northwestern University to the first author. We are grateful to the toddlers and caretakers who participated in these studies. We are also indebted to Molly Niendorf, Gwen Fiske and several other students for their assistance with data collection and coding.

## REFERENCES

- Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology*, *14*(1), 79–93.
- Akhtar, N., & Tomasello, M. (2000). The social nature of words and their meaning. In R. M. Golinkoff, K. Hirsh-Pasek, et al. (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). New York: Oxford University Press.

- Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20(2), 395–418.
- Baldwin, D. A. (Ed.). (1995). *Understanding the link between joint attention and language*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bauer, P. J. (1992). Holding it all together: How enabling relations facilitate young children's event recall. *Cognitive Development*, 7(1), 1–28.
- Booth, A. E., & Waxman, S. R. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology*, 38(6), 948–957.
- Bradshaw, G. L., & Anderson, J. R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning & Verbal Behavior*, 21(2), 165–174.
- Brand, R. J. (2000). Learning novel nouns: Children use multiple cues. *Monographs of the Society for Research in Child Development*, 65(3), 41–61.
- Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': Modifications in mothers' infant-directed action. *Developmental Science*, 5(1), 72–83.
- Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8(6), 535–543.
- Butler, S. C., Caron, A. J., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition & Development*, 1(4), 359–377.
- Butterworth, G., & Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology*, 18(1), 25–50.
- Campbell, A. L., & Namy, L. L. (2003). The role of social-referential context in verbal and nonverbal symbol learning. *Child Development*, 74(2) Mar-Apr 2003, 549–563.
- Capone, N. C., & McGregor, K. K. (2005). The effect of semantic representation of toddlers' word retrieval. *Journal of Speech, Language, & Hearing Research*, 48, 1468–1480.
- Caron, A. J., Kiel, E. J., Dayton, M., & Butler, S. C. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. *Journal of Cognition & Development*, 3(4), 445–464.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), 176.
- Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, 7(4), 303–319.
- Copple, C. E., & Coon, R. C. (1977). The role of causality in encoding and remembering events as a function of age. *Journal of Genetic Psychology*, 130(1), 129–136.
- Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, 34(1), 28–38.
- Dannemiller, J. L., & Freedland, R. L. (1993). Motion-based detection by 14-week-old infants. *Vision Research*, 33(5–6), 657–664.
- Elman, J. L. (2001). Connectionism and language acquisition. In M. Tomasello & E. Bates (Eds.), *Language development: The essential readings* (pp. 295–306). Malden, MA: Blackwell Publishers.
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J., et al. (1993). *User's guide and technical manual for the MacArthur communicative development inventories*. San Diego, CA: Singular Press.
- Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, 58(6), 1532–1541.
- Gogate, L. J., & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology*, 69(2), 133–149.

- Gogate, L. J., & Bahrick, L. E. (2001). Intersensory redundancy and 7-month-old infants' memory for arbitrary syllable-object relations. *Infancy*, 2(2), 219–231.
- Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, 71(4), 878–894.
- Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). The intersensory origins of word comprehension: An ecological-dynamic systems view. *Developmental Science Vol 4(1)*, 1–18.
- Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development*, 75(2), 409–427.
- Hennon, E. (2000). What does it take for 12-month-olds to learn a word? *Monographs of the Society for Research in Child Development*, 65(3), 62–84.
- Hirsh-Pasek, K., Golinkoff, M., & Hollich, G. (2000). An emergentist coalition model for word learning: Mapping words to objects is a product of the interaction of multiple cues. In R. M. Golinkoff, K. Hirsh-Pasek, et al. (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 136–164). New York: Oxford University Press.
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., et al. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65(3), v–123.
- Horst, J. S., & Samuelson, L. (in press). Fast mapping but poor retention by 24-month-old infants. *Infancy*.
- Kellman, P. J. (1993). Kinematic foundations of infant visual perception. In C. Granrud (Ed.), *Visual perception and cognition in infancy. Carnegie Mellon symposia on cognition* (pp. 121–173). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science Vol 7(3)*, 297–307.
- Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, 7(2), 173–187.
- McGregor, K. K., Newman, R. M., Reilly, R., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, & Hearing Research*, 45(5), 998–1014.
- Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental Psychology*, 35(1), 60–68.
- Morisette, P., Ricard, M., & Gouin Decarie, T. (1995). Interactions between shared attention during the preverbal period and the development of verbal references. *International Journal of Psychology*, 30(4), 481–498.
- Namy, L. L. (2001). What's in a name when it isn't a word? 17-month-olds' mapping of nonverbal symbols to object categories. *Infancy*, 2(1), 73–86.
- Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development*, 77(2), 266–280.
- Rovee-Collier, C., & Hayne, H. (2000). Memory in infancy and early childhood. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 267–282).
- Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, 69(1), 94–104.
- Smith, E. E., & Kosslyn, S. M. (2007). *Cognitive psychology: Mind and brain*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In C. A. Nelson (Ed.), *Basic and applied perspectives on learning, cognition, and development. The Minnesota Symposia on Child Psychology* (Vol. 28, pp. 1–32). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, 60(2), 143–171.

- Thoermer, C., & Sodian, B. (2001). Preverbal infants' understanding of referential gestures. *First Language, Vol 21(63)*, 245–264.
- Tomasello, M. (1992). The social bases of language acquisition. *Social Development, 1(1)*, 67–87.
- Tomasello, M. (1995). Joint attention as social cognition. In C. E. D. Moore, J. Philip (Eds.), *Joint attention: Its origins and role in development* (pp. 104–130). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc, Publishers.
- Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In E. Bates & M. Tomasello (Eds.), *Language development: Essential readings* (pp. 111–128). Malden, MA: Blackwell Publishers Inc.
- Tomasello, M. (2003). The key is social cognition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 47–57). Cambridge, MA: MIT Press.
- Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology, 30(5)*, 639–650.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78(3)*, 705–722.
- Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language, 23(1)*, 157–176.
- Valenzano, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28(2)*, 187–204.
- Waxman, S. R., & Booth, A. E. (2000). Principles that are invoked in the acquisition of words, but not facts. *Cognition, 77(2)*, B33-B43.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34(6)*, 1289–1309.
- Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology, 30(4)*, 553–566.
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69(1)*, 134.
- Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science, 6(3)*, 297–311.
- Woodward, A. L. (2004). Infants' Use of Action Knowledge to Get a Grasp on Words. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 149–172). Cambridge, MA: MIT Press.
- Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development, 17(1)*, 1061–1084.
- Zukow-Goldring, P. (2006). Assisted imitation: Affordances, effectivities, and the mirror system in early language development. In M. A. Arbib (Ed.), *From action to language via the mirror neuron system* (pp. 469–500). Cambridge, England: Cambridge University Press.

APPENDIX

<i>Activity</i>	<i>Procedure</i>
Warm up	E and C* play with familiar toys for approximately 5 minutes
Word teaching episode 1	<ol style="list-style-type: none"> <li>1. E displays target-foil pair 1 and produces nonce word three times while using condition-specific gestures to refer to target</li> <li>2. C plays with target-foil pair 1 for 10s.</li> <li>3. E again displays target-foil pair 1 and produces nonce word three times while using condition-specific gestures to refer to target</li> <li>4. Steps 1 – 3 are repeated with target-foil pairs 2 and 3**</li> </ol>
Choice task training	<ol style="list-style-type: none"> <li>1. C plays with set of training objects for 30s.</li> <li>2. E arrays objects on tray, introduces dog puppet, and “Doggy” requests spoons (or books)</li> <li>3. C responds by handing object to Doggy</li> <li>4. E praises or corrects response</li> <li>5. E asks, “Are there any more spoons (or books) for Doggy?”</li> <li>6. C responds</li> <li>7. E praises or corrects response</li> <li>8. Steps 1 – 7 are repeated with a second training set</li> </ol>
Word teaching episode 2	<ol style="list-style-type: none"> <li>1. E displays target-foil pair 1 and produces nonce word three times while using condition-specific gestures to refer to target</li> <li>2. C plays with target-foil pair 1 for 10s.</li> <li>3. E again displays target-foil pair 1 and produces nonce word three times while using condition-specific gestures to refer to target</li> </ol>
Production testing	<ol style="list-style-type: none"> <li>1. E displays target-foil pair 1; introduces bunny puppet; “Bunny” requests name of target and name of foil</li> <li>2. C responds</li> <li>3. E provides neutral feedback (“ok”)</li> </ol>
Comprehension testing	<ol style="list-style-type: none"> <li>1. C plays with set 1 comprehension test objects for 30s.</li> <li>2. E arrays test objects in random order on tray; “Bunny” requests novel target 1 (e.g., koob)</li> <li>3. C responds</li> <li>4. E asks “are there any more koobs for Bunny?” up to three times or until C responds “no”</li> </ol>
Word teaching episode 2; Production and Comprehension testing	These 3 activities, described above, are repeated with target-foil pair 2
Word teaching episode 2; Production and Comprehension testing	These 3 activities, described above, are repeated with target-foil pair 3
Follow-up testing	Production and comprehension testing repeated

\*E = examiner; C = child.

\*\*order of presentation of object sets 1 – 3 varied randomly across children.