

# Clarifying the role of joint attention in early word learning

First Language

1–16

© The Author(s) 2011

Reprints and permission: [sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)

DOI: 10.1177/0142723710395423

[fla.sagepub.com](http://fla.sagepub.com)



**Jason Scofield**

University of Alabama, USA

**Douglas A. Behrend**

University of Arkansas, USA

## Abstract

Four studies examined whether two-year-olds could successfully learn a novel word in conditions in which joint attention was not present. Study 1 examined whether children could learn a novel word while the speaker, but not the child, attended to the target object. Study 2 examined whether children could learn a novel word while the child, but not the speaker, attended to the target object. Study 3 examined whether children could learn a novel word while the child and the speaker attended to two different target objects. Study 4 examined whether children could learn a novel word while neither the child nor the speaker attended to the target object. Findings showed that successful word learning occurred in each of the four studies. These results suggest that joint attention may play an important, though not a necessary, role in young children's word learning.

## Keywords

children, joint attention, language development, referential intent, word learning

It is generally agreed that, when present, joint attention is important for successful word learning (Akhtar, 2005a; Baldwin, 1995; Tomasello, 1995). However, there is less agreement on the likelihood of successful word learning when joint attention is not present.

An episode of joint attention occurs when partners attend to the same target (i.e., joint focus) and understand the nature of their shared attention (i.e., intersubjective awareness) (Akhtar, 2005a; Baldwin, 1995). At least two important inferences are believed to be made during joint attention. First, joint attention is inferred to be a referential

---

## Corresponding author:

Jason Scofield, 222 Child Development Research Center, Human Development and Family Studies, College of Human Environmental Sciences, University of Alabama, Tuscaloosa, AL 35487-0160, USA.

Email: [scofield@ches.ua.edu](mailto:scofield@ches.ua.edu)

framework that includes both the attending partners and the shared target (Akhtar, 2005a; Bakeman & Adamson, 1984; Baldwin, 1995; Baldwin et al., 1996; Baldwin & Moses, 2001; Bruner, 1999; Moore, Angelopoulos, & Bennett, 1999; Tomasello, 1995). Second, the shared target is inferred to be the intended referential target (Baldwin, 1991, 1993a, 1995; Baldwin et al., 1996; see also Sabbagh & Baldwin, 2005; Tomasello, 1995, 2001).

These inferences are important because they help resolve two problems that are common for children learning words. First, children must distinguish those acts that are intended to refer from those acts that are not intended to refer (i.e., referential vs. non-referential acts) (Baldwin, 1993a; Baldwin et al., 1996; Moore et al., 1999). Second, children must distinguish those targets that are intended to be referred to from other potential targets (i.e., referential vs. non-referential targets) (Baldwin, 1993a; Baldwin & Tomasello, 1998; Behrend, 1990; Markman, 1990; Quine, 1960). Joint attention helps resolve both of these problems. As a shared referential state, joint attention readies the child for a referential act and then identifies the likely referential target.

It thus could be said that joint attention provides some of the most optimal conditions for successful word learning. During joint attention, the child: (1) attends to the shared target with the speaker, (2) understands the speaker's reference is intentional, and (3) concludes that the speaker's intent is to refer to the shared target. Under these conditions the task of word learning appears to be at its most transparent and thus most likely to be successful (Moore et al., 1999). It is therefore not surprising that joint attention has been argued to play a very important, and even a necessary, role in successful word learning (Akhtar, 2005a; Akhtar, Jipson, & Callanan, 2001; Akhtar & Tomasello, 1998; Baldwin, 1995; Tomasello, 1995).

Support for this important role comes from previous studies that have shown that the presence of joint attention promotes successful word learning. For example, early episodes of joint attention in infancy are related to later word learning success (Carpenter, Nagell, & Tomasello, 1998; Markus, Mundy, Morales, Delgado, & Yale, 2000; Saxon, 1997; Smith, Adamson, & Bakeman, 1988; Tomasello & Farrar, 1986; see Slaughter & McConnell, 2003 for contrary findings). Also, when establishing joint attention with infants, attention-following results in greater language gains than attention-switching or attention-directing (Akhtar, Dunham, & Dunham, 1991; Dunham, Dunham, & Curwin, 1993; Tomasello & Farrar, 1986; Tomasello & Todd, 1983). In addition, when joint attention with a social partner is disrupted toddlers show an increase in referential behavior including more pointing and more vocalizations (Franco & Gagliano, 2001). Finally, upon hearing a novel word, toddlers regularly consult the speaker for referential cues as to the referent of that word (Baldwin, 1991, 1993a, 1993b; see also Sabbagh & Baldwin, 2005). These studies provide evidence that the presence of joint attention facilitates successful word learning.

Interestingly however, not all previous studies have shown that joint attention is a prerequisite for word learning. For example, successful word learning has been shown in conditions in which the speaker is absent (Rice, Huston, Truglio, & Wright, 1990; Rice & Woodsmall, 1988; Schafer & Plunkett, 1998; Scofield, Williams, & Behrend, 2007), the speaker makes an ambiguous reference (e.g., Akhtar & Montague, 1999; Carey & Bartlett, 1978; Markman & Wachtel, 1988; Merriman & Bowman, 1989), joint focus is

absent (Akhtar & Tomasello, 1996; Baldwin, 1991, 1993b; Tomasello & Kruger, 1992), and intersubjective awareness is absent (Akhtar, 2005b; Akhtar et al., 2001; Jaswal & Markman, 2003). Importantly, though joint attention was not present in these examples (at least in its optimal state), the overall impact on the success of word learning was small. These studies suggest that the absence of joint attention does not prevent successful word learning and, consequently, make unclear the nature of the role of joint attention in word learning.

The Akhtar et al. (2001) studies offer an especially interesting departure point for a discussion on joint attention's role in word learning (see also Floor & Akhtar, 2006). In Akhtar et al., two-year-olds were presented with an addressed condition or an overhearing condition. In the addressed condition the child directly participated in an interaction with a speaker, while in the overhearing condition the child observed an interaction between a speaker and another adult. In the addressed condition the child and speaker retrieved four unfamiliar objects from a box, one of which was labeled with a novel word and returned to the box and three of which were returned unlabeled. The overhearing condition was similar except that an assistant and the speaker retrieved and returned the objects while the child was positioned as an onlooker. Akhtar et al. showed that successful word learning did not differ across the two conditions and ultimately suggested that intersubjective awareness, previously argued to play a fundamental role in joint attention (Baldwin, 1995; see also Akhtar, 2005b; Jaswal & Markman, 2003), may not be important for word learning.

There are at least two possible explanations for successful word learning in Akhtar et al.'s (2001) study. One explanation is that the presence of intersubjective awareness is not necessary for joint attention (see Akhtar, 2005a for further discussion). However, there is at least some reason to be skeptical of this explanation given what would remain of an episode of joint attention without intersubjective awareness. Joint attention would then consist simply of the coordinated attention between both the child and the speaker. Baldwin (1995; see also Baldwin et al., 1996; Tomasello, 1995; Tomasello & Farrar, 1986) has argued convincingly that simply attaining a coordinated attentional state may not allow the child to understand the referential nature of joint attention.

A second explanation is that presence of joint attention is not necessary for successful word learning. On the surface this might seem like a surprising explanation given the focus placed on joint attention in the language development literature for the past three decades. However, considering the findings of the studies already reviewed, the challenge of establishing joint attention before learning each new word (see Akhtar, 2005a for discussion), and other factors like the vocabulary growth experienced by some children with joint attention deficits (e.g., children with autism spectrum disorders) (see Akhtar & Gernsbacher, 2007 for discussion), this may be the more viable of the two explanations. Though, to date there has not been a systematic examination of word learning in the absence of joint attention.

The current studies conducted this examination. If word learning depends critically on joint attention then word learning should significantly decline, or be altogether unsuccessful, in the absence of joint attention. Conversely, if word learning does not depend critically on joint attention then word learning should not be affected by the absence of joint attention.

## Method

Four separate studies were conducted ( $N = 48$  in each study). In each study a child was presented with a word while important components of joint attention were missing. In Study 1 the novel word was presented while the speaker, but not the child, attended to a target object. In Study 2 the novel word was presented while the child, but not the speaker, attended to a target object. In Study 3 the novel word was presented while the child and the speaker attended to two different target objects (see Baldwin, 1991). In Study 4 the novel word was presented while neither the child nor the speaker attended to a target object. If the presence of joint attention is needed for successful word learning then word learning should be unsuccessful in these studies.

## Study 1

### *Participants*

Forty-eight two-year-olds, 21 girls and 27 boys ( $M = 2;7$ ,  $SD = 0;3$ ), completed Study 1. Thirty-two of the children ( $M = 2;7$ ,  $SD = 0;3$ ) were assigned to the experimental condition while the remaining 16 were assigned to the control condition ( $M = 2;7$ ,  $SD = 0;3$ ). Parent consent and child assent were obtained from participating children enrolled at preschools near a university campus in the southwestern United States.

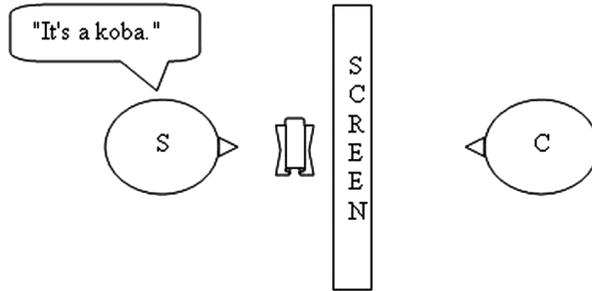
### *Materials*

A wooden partition was used to separate the child from the speaker in both the experimental and control conditions. The partition was constructed of two wooden panels, a screen, and a base. The screen measured 20 inches long  $\times$  15 inches tall  $\times$  1 inch wide while the base measured 20 inches long  $\times$  6 inches wide  $\times$  1 inch thick. The screen was fixed perpendicularly to the base allowing the partition to be self-standing and resulting in two small ledges on which the target object could be positioned. The partition was designed to obstruct attention to a target but not to obstruct attention to a partner.

The target object was selected from a set of four unfamiliar objects which included a plastic T-joint (measuring 3 inches  $\times$  2 inches  $\times$  1 inch), a metal shelving bracket (measuring 2 inches  $\times$  2 inches  $\times$  1 inch), a plastic elbow joint (measuring 2 inches  $\times$  1 inch  $\times$  1 inch), and a plastic mounting bracket (measuring 2 inches  $\times$  2 inches). A set of four unfamiliar distracters was also selected. Objects were part of separate and discrete sets and previous testing using these objects revealed that they were unfamiliar to two-year-olds (see Behrend, Scofield, & Kleinknecht, 2001).

### *Procedure*

The child and the speaker sat opposite one another at a small table where the child completed a warm-up and a testing session. The warm-up session was designed to familiarize the child with both the speaker and the testing procedure. During the warm-up the child was presented with a set of wooden building blocks and asked to sort the blocks by select properties (e.g., 'Can you find the blocks with blue letters?' or 'Which blocks have trains



**Figure 1.** Overhead view of Study 1 procedure: Speaker attends experimental condition

on them?’). After two to four trials (depending on participation) the blocks were removed and the testing session began.

In both the experimental condition and the control condition the partition was positioned between the child and speaker with a target object already in view of the speaker. The partition obstructed the child’s view of the target, but not the child’s view of the speaker or vice versa (see Figure 1).

In the experimental condition, the speaker presented a novel word three times while alternating eye contact between the child and the target (e.g., ‘Wow, a koba. A koba! It’s a koba.’). Following a 5 second delay, the speaker moved the partition so that the target was visible to both the speaker and the child. The target was then removed completely from view and included in an array with four distracters. The array was then placed on the table and the child was asked to select the object that best corresponded to the novel word (e.g., ‘Can you help me find the koba? I can’t find the koba. Which one of these is the koba?’). The speaker was careful not to provide any non-verbal cues during object selection.

The control condition was identical to the experimental condition except that a neutral comment, rather than a novel word, was presented three times (e.g., ‘It’s fun. Wow! It’s neat.’). The control condition was designed to account for any effect that exposure had on object selection. The target, distracters, and words used in the control condition were selected randomly and differed from those in the experimental condition.

## Results – Study 1

Of interest in Study 1 was whether word learning would be successful when the child did not attend to a target at the time of labeling. The dependent measure was the child’s object selection, either the target or a distracter, with chance performance in both the experimental and control conditions being equal to 20%. However, in the experimental condition 72% of children selected the target (i.e., 23 of 32) (see Table 1). Chi-square analysis revealed that this proportion was significantly different from chance,  $\chi^2(1, N = 32) = 53.82, p < .01$ . In the control condition, only 19% of children selected the target (i.e., 3 of 16). Chi-square analysis also revealed that children selected the target significantly more often in the experimental condition than in the control condition,  $\chi^2(1, N = 48) = 12.13, p < .01$ .

**Table 1.** Percentage (and number) of children selecting the target(s) in Studies 1 to 4

Study	Experimental condition	Control condition	Chance
Study 1 (Speaker attends)	72% (23/32) <sup>a,b</sup>	19% (3/16)	20%
Study 2 (Child attends)	72% (23/32) <sup>a,b</sup>	6% (1/16)	20%
Study 3 (Both attend)	75% (18/24) <sup>a,b</sup>	25% (6/24)	40%
Study 4 (Neither attends)	46% (11/24) <sup>a</sup>	21% (5/24)	20%

<sup>a</sup>Different from chance,  $p < .05$ .

<sup>b</sup>Different from control condition,  $p < .05$ .

## Discussion – Study 1

The results of Study 1 suggest that the absence of the child's attention during labeling does not lead to unsuccessful word learning. In Study 1 children successfully mapped a novel word to a novel object despite not having attended to the object when the word was being presented. These results are impressive considering that successful word learning required children to anticipate a potential referent for the word and then wait until that referent appeared, something altogether different from jointly attending to a shared referent. However, while these results are compelling, they may not alone have definitive implications for the role of joint attention in successful word learning. As stated earlier, joint attention requires the convergence of multiple cues – including the child's attention. It is unclear however whether these cues are equally important for the task of word learning or whether one cue may be more or less important than another (see Akhtar et al., 2001 for discussion of intersubjective awareness). For example, the speaker's attention has been described as an especially important cue because it indicates referential intent (Akhtar, 2005a; Baldwin, 1991, 1993a, 1993b; Baldwin et al., 1996; see also Akhtar et al., 2001).

In Study 1, the speaker's attention was focused on the target object, which may have led the child to prefer that target as the intended referent. While this explanation still implies that the child did not actually attend to the target, the child nonetheless found some reason to prefer the target over the distracters and this preference was not explained by exposure. To better determine the significance of the speaker's attention, in Study 2 the novel word is presented while the child, but not the speaker, attends to the target.

## Study 2

### Participants

Forty-eight two-year-olds, 20 girls and 28 boys ( $M = 2;6$ ,  $SD = 0;3$ ), completed Study 2. The participants in Study 2 were different from the participants in Study 1. Thirty-two of the children ( $M = 2;6$ ,  $SD = 0;3$ ) were assigned to the experimental condition while the remaining 16 were assigned to the control condition ( $M = 2;7$ ,  $SD = 0;4$ ). Parent consent and child assent were obtained from participating children enrolled at preschools near a university campus in the southwestern United States.

## Materials

The materials used in Study 2 were identical to the materials used in Study 1.

## Procedure

The procedure in Study 2 was very similar to the procedure in Study 1 except that the target was positioned in view of the child such that the partition obstructed the speaker's view of the target, but not the speaker's view of the child or vice versa. Also, the word and comment were presented while the speaker alternated eye contact between the child and a designated spot on the screen (marked 'x') (e.g., 'Wow, a nixon. A nixon! It's a nixon. '), rather than between the child and the target. The target, distracters, and words used differed across condition and from those used in Study 1.

## Results – Study 2

Of interest in Study 2 was whether word learning would be successful when the speaker did not attend to a target. Indeed, in the experimental condition 72% of children selected the target (i.e., 23 of 32) (see Table 1). Chi-square analysis revealed that this proportion was significantly different from chance,  $\chi^2(1, N = 32) = 53.82, p < .01$ . In the control condition, only 6% of children selected the target (i.e., 1 of 16). Chi-square analysis also revealed that children selected the target significantly more often in the experimental condition than in the control condition,  $\chi^2(1, N = 48) = 11.15, p < .01$ .

## Discussion – Study 2

The results of Study 2 suggest that the absence of the speaker's attention does not lead to unsuccessful word learning. In Study 2 children successfully mapped a novel word to a novel object despite the speaker not having attended to the object when the word was presented. Again though, while these results are certainly compelling and address an issue that remained from Study 1, they may also not have definitive implications for the role of joint attention in successful word learning. In Study 2, the speaker's attention was not focused on a target object – something the child could not have known. It is possible that the child assumed the speaker was attending to a target and preferred that target as the intended referent. When no such target was revealed, the child may have simply defaulted to the next logical selection – the object they did attend to. While this explanation still implies that the speaker did not actually attend to the target, the child nonetheless found some reason to prefer the target over the distracters and this preference was not explained by exposure. To better determine the child's target preference, in Study 3 the novel word is presented while the child and speaker attend to two different targets.

## Study 3

### Participants

Forty-eight two-year-olds, 28 girls and 20 boys ( $M = 2;7, SD = 0;5$ ), completed Study 3. The participants in Study 3 were different from the participants in Study 1 and Study 2.

Twenty-four of the children ( $M = 2;7$ ,  $SD = 0;4$ ) were assigned to the experimental condition while the remaining 24 were assigned to the control condition ( $M = 2;7$ ,  $SD = 0;5$ ). Parent consent and child assent were obtained from participating children enrolled at preschools near a university campus in the southeastern United States.

## Materials

The materials used in Study 3 were identical to the materials used in Study 1 and Study 2.

## Procedure

The procedure in Study 3 was very similar to the procedures in Studies 1 and 2 except that the child and the speaker each viewed a different target. The partition obstructed the child's view of the speaker's target and the speaker's view of the child's target, but not the child and speaker's view of one another. Also, the word and comment were presented while the speaker alternated eye contact between the child and the speaker's target (e.g., 'There's the jeter. A jeter! It's a jeter.'). Following a 5 second delay, the speaker moved the partition so that both targets were visible to both the speaker and the child. The targets were then removed completely from view and included in an array with three distracters. The target, distracters, and words used differed across condition and from those used in Studies 1 and 2.

## Results – Study 3

Of interest in Study 3 was whether word learning would be successful when the child and the speaker attended to two different targets. Chance was equal to 40% as there were two possible targets. However, in the experimental condition 75% of children selected one of the targets (i.e., 18 of 24) (see Table 1). Chi-square analysis revealed that this proportion was significantly different from chance,  $\chi^2(1, N = 24) = 12.25$ ,  $p < .01$ . In the control condition, only 25% of children (i.e., 6 of 24) selected one of the targets. Chi-square analysis also revealed that children selected the target significantly more often in the experimental condition than in the control condition,  $\chi^2(1, N = 48) = 12.00$ ,  $p < .01$ .

Of additional interest was which of the two targets, the child's or the speaker's, would be selected as the target. Chance for each target was equal to 20%. However, while 25% of children selected the speaker's target (i.e., 6 of 24), 50% of children selected the child's target (i.e., 12 of 24) (see Table 2). Chi-square analysis revealed that selection of the child's target differed significantly from chance,  $\chi^2(1, N = 24) = 13.50$ ,  $p < .01$ . In the control condition, only 13% of children selected the speaker's target (i.e., 3 of 24) and only 13% of children selected the child's target (i.e., 3 of 24). Chi-square analysis also revealed that selection of the child's target different significantly from selection of the speaker's target,  $\chi^2(1, N = 24) = 8.00$ ,  $p < .01$ . In contrast to previous findings (Baldwin, 1991, 1993b), these results suggest that children do not always prefer the target of the speaker's attention.

**Table 2.** Percentage (and number) of children selecting the speaker-attended vs. the child-attended target in Study 3

Target	Experimental condition	Control condition	Chance
Speaker attended	25% (6/24)	13% (3/24)	20%
Child attended	50% (12/24) <sup>a,b</sup>	13% (3/24)	20%

<sup>a</sup>Different from chance,  $p < .05$ .

<sup>b</sup>Different from control condition,  $p < .05$ .

### Discussion – Study 3

The results of Study 3 suggest that successful word learning can occur when the child and speaker attend to different objects. Further, the results suggest that, when attentions diverge, the child may actually prefer the target of their own attention. There are two interesting points to make about these results. First, even in an ambiguous setting the child's preference was to select a target and not a distracter. Second, there was not a preference to select the speaker's target, a potential explanation for the results in Studies 1 and 2.

Study 3 is not the first to include a child and speaker with divergent, or discrepant, attentions. Well-known studies conducted by Baldwin (1991, 1993b) used a similar design – though there are notable differences between the studies including age, analysis, and results. Children in Study 3 were older (i.e., 2;7) than children in the Baldwin studies (i.e., 1;6). In Study 3, selection of the child's target and selection of the speaker's target were treated as a single value and compared to chance whereas in the Baldwin studies selection of the child's target was compared to the selection of the speaker's target. And finally, in Study 3 children preferred the target of their own attention, which contrasts with Baldwin's reports that the child's target was less likely or unlikely to be selected. It is possible that the first two differences between the studies account for the third.

Thus far the results of Studies 1 to 3 straightforwardly suggest that the success of word learning does not depend on the presence of joint attention. However, there may be reason to believe that word learning would be affected in the absence of attention alone. That is, word learning could be affected if no one attended to the target. In each of the previous studies the child, the speaker, or both have attended to a target and so do not speak to the importance of attention alone. To better determine the role of attention in word learning, the novel word in Study 4 is presented while neither the child nor speaker attended to the target.

## Study 4

### Participants

Forty-eight two-year-olds, 30 girls and 18 boys ( $M = 2;8$ ,  $SD = 0;6$ ), completed Study 4. The participants in Study 4 were different from the participants in Studies 1, 2, and 3. Twenty-four of the children ( $M = 2;9$ ,  $SD = 0;6$ ) were assigned to the experimental condition while the remaining 24 were assigned to the control condition ( $M = 2;6$ ,  $SD = 0;4$ ).

Parent consent and child assent were obtained from participating children enrolled at preschools near a university campus in the southeastern United States.

### *Materials*

A wooden box was used to obstruct attention to the target in both the experimental and control conditions. Each side of the box measured 8 inches making the dimensions of the box 8 inches long  $\times$  8 inches tall  $\times$  8 inches wide, and included a hinged top. The box was designed to obstruct attention to a target but not to obstruct attention to a partner. The object sets in Study 4 were like those in Studies 1 to 3.

### *Procedure*

The procedure in Study 4 was very similar to the procedure in Studies 1 to 3 except that the target was concealed in a box such that the box obstructed both the child and speaker's view of the target, but not the child's view of the speaker or vice versa. Also, the word and comment were presented while the speaker alternated eye contact between the child and a designated spot on the top of the box (marked 'x') (e.g., 'Wow, an agnew. An agnew! It's an agnew.'). not unlike Study 2. Finally, following a 5 second delay, the speaker removed the target from the box and placed the target on the table so that it was visible to both the speaker and the child. The target was then removed completely from view and included in an array with four distracters. The target, distracters, and words used differed across condition and from those used in Studies 1 to 3.

### *Results – Study 4*

Of interest in Study 4 was whether word learning would be successful when neither the child nor the speaker attended to a target. Chance was equal to 20%. However, in the experimental condition 45% of children selected the target (i.e., 11 of 24) (see Table 1). Chi-square analysis revealed that this proportion was significantly different from chance,  $\chi^2(1, N = 24) = 10.01, p < .01$ . In the control condition, only 21% of children selected the target (i.e., 5 of 24). Chi-square analysis also revealed that children selected the target marginally more often in the experimental condition than in the control condition,  $\chi^2(1, N = 48) = 3.375, p = .062$ . Consequently, Study 4 does not exclude the possibility that children's selection of the target was based on exposure.

### *Discussion – Study 4*

The results of Study 4 suggest that successful word learning can occur when neither the child nor the speaker attend to a target object. However, the overall success in Study 4 (45%) was noticeably lower than in Studies 1, 2, or 3 (72–75%) suggesting that word learning may have been negatively affected by the absence of attention to the target. Also, while word learning in the experimental condition did differ significantly from chance it did not differ from word learning in the control condition. This suggests that target exposure could not be dismissed as a possible explanation for target selection.

Overall, these results suggest that the success of word learning may be inhibited, but likely not precluded, when neither the speaker nor the child attends to a target.

## General discussion

In order to successfully learn a new word the child must understand the speaker's referential intent (Akhtar, 2005a; Baldwin, 1993a; Baldwin, 1995; Baldwin et al., 1996; Baldwin & Moses, 2001; Bruner, 1999; Moore et al., 1999; Sabbagh & Baldwin, 2005; Tomasello, 1995, 2001) and must correctly identify the intended referent (Baldwin, 1991, 1993a, 1995; Baldwin et al., 1996). Joint attention is believed to assist in accomplishing both, thereby creating what are often considered the optimal conditions for learning a word. As a result, joint attention has been described as a necessary part of successful word learning (Akhtar, 2005a; Akhtar et al., 2001; Akhtar & Tomasello, 1998; Baldwin, 1995; Tomasello, 1995). However, there are now several examples of successful word learning under non-optimal conditions of joint attention – including the current studies – that challenge this description. These examples raise at least two important questions about the relationship between joint attention and word learning. One, what is the role of joint attention in successful word learning? And two, what are the necessary conditions for successful word learning?

### *What is the role of joint attention in successful word learning?*

When present, joint attention plays a clear and important role in successful word learning – decades of research shows that. The question here however was whether successful word learning could occur when joint attention was not present. The results of the current studies show that it can. Indeed the rate of word learning in the absence of joint attention in these studies did not appear to differ from the rate of word learning reported in previous studies where joint attention was present (Behrend et al., 2001; Krcmar, Grela, & Lin, 2007). Moreover, the similarities between the rates of word learning in the presence and absence of joint attention were not unlike the similarities between rates of word learning in referent-absent and referent-present conditions (Akhtar & Tomasello, 1996), in speaker-absent and speaker-present conditions (Krcmar et al., 2007; Scofield et al., 2007), or in conditions in which intersubjective awareness is absent or present (Akhtar, 2005a, 2005b; Akhtar et al., 2001). Ultimately, the absence of joint attention, or important components of joint attention, does not appear to negatively impact children's ability to learn a word – suggesting that joint attention does not play a necessary role in young children's word learning (Akhtar, 2005b; Akhtar & Gernsbacher, 2007; Moore et al., 1999).

Left unclear though is the specific role that joint attention does play in word learning, though there are several possibilities. It is possible that joint attention is critical for word learning, but only for children younger than two years old. Baldwin (e.g., 1993a, 1993b) and others have shown that children younger than two years old may use joint attention differently than their older counterparts. It is possible that joint attention plays a leading role in word learning when present, but does not negatively impact word learning when absent. It is possible that joint attention combines with other social and language cues to

facilitate word learning. It is possible that other cues would facilitate a child's initial hypothesis about the meaning of a word and joint attention simply helps to confirm that hypothesis. Each of these possibilities allows joint attention to play an important role in the word learning process when present and yet not preclude word learning when absent. Though the current studies do not differentiate between these various roles, they set the stage for future studies to do so.

Finally, it could be argued that understanding a speaker's referential intent is the critical factor for word learning and that joint attention reveals referential intent (Akhtar, 2005a; Baldwin, 1993a, 1995; Baldwin et al., 1996; Baldwin & Moses, 2001; Bruner, 1999; Moore et al., 1999; Sabbagh & Baldwin, 2005; Tomasello, 1995, 2001). Consequently, it is possible that joint attention is not critical for word learning, but that it serves a critical role by revealing a speaker's intent.

However, the current studies do not support this argument. In these studies, performance in the control conditions differed dramatically from performance in the experimental conditions. This was true even though the speaker's actions did not differ across conditions. In fact, the only difference across conditions was the presence or absence of the novel word. Therefore, if children are inferring referential intent in the experimental condition (where they learned words) but not in the control condition (where they did not learn words) then the implication is that the word, not the speaker, is supporting the inference.

In addition, despite using different procedures, each of the current studies led to word learning. Using referential intent to explain word learning in each study would require the child to make at least three separate inferences. The child would have to infer that the speaker intended to refer to: (a) the speaker's target in Study 1, (b) the child's target in Study 2, (c) either the speaker's or the child's target in Study 3, and (d) a target that no one had seen in Study 4. Rather than using different inferences to explain word learning across these studies (and across conditions), the simpler explanation might be that the child was using something other than referential intent when deciding the referent of the word. Therefore, even if joint attention's role in word learning is to reveal referential intent, the current results suggest that this role is probably not crucial for learning a word.

### *What are the necessary conditions for successful word learning?*

As outlined earlier, during joint attention the child and speaker attend to the same target and understand the shared nature of their attention. In addition, the child understands that the speaker's intentions are referential and directed at the target (see Figure 2). Using the results of current and previous studies it may be possible to isolate which, if any, of these conditions seem necessary for successful word learning. For example, Studies 1 and 2 suggest that word learning does not depend on the child or the speaker attending to the target. Akhtar et al.'s (2001) overhearing conditions suggest that word learning does not depend on intersubjective awareness. Scofield et al.'s (2007) speaker-absent conditions suggest that word learning does not depend on the presence of a speaker.

Removing these components from the model depicted in Figure 2 leaves only three remaining components that word learning might depend on: (1) a word, (2) someone to learn the word (i.e., the child), and (3) something to learn the word about (i.e., the referent) – although the referent need not be present during labeling. According to this

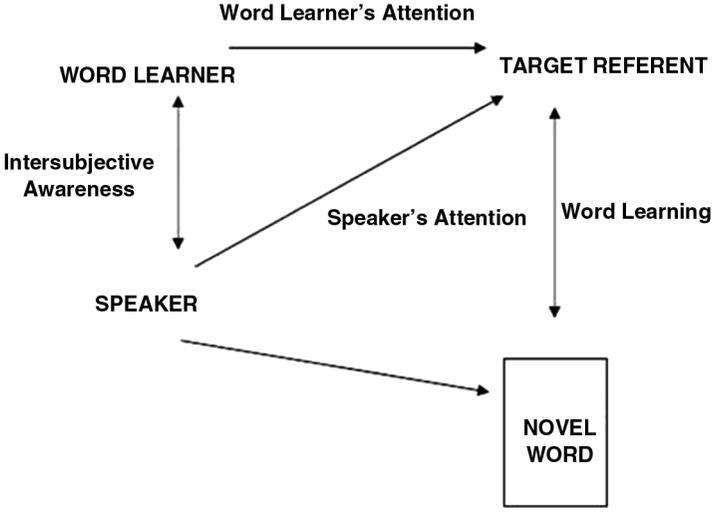


Figure 2. A conceptual model of word learning during joint attention

thinking, these remaining components would make up the minimal conditions for word learning and the other components, while facilitative, would not be necessary. Although, it should be noted that it is not clear from the current studies or from previous studies that word learning would actually succeed if only the word, child, and referent were present (though see Scofield et al., 2007).

It is also not clear from these studies whether the conditions that might support learning a novel noun would also support learning a novel verb or adjective. Likewise, it is not clear from these studies whether the conditions that might support learning a novel noun in two-year-olds would also support learning a novel noun in either older or younger children. The same could be said for the role of joint attention in word learning. That is, these studies do not clarify the role of joint attention in learning other types of words (e.g., verbs or adjectives) or in word learning in younger or older children. Instead, these studies clarify the role that joint attention has in young children’s early word learning and suggest that joint attention is important, but not critical, for two-year-olds learning words.

### Conclusions

These studies examined the role of joint attention in young children’s word learning. Building from previous work (Akhtar, 2005a; Akhtar et al., 2001), the current studies are among the first designed to systematically test the availability of joint attention and its resulting effect on word learning and likewise are among the first to explicitly find that word learning occurs in the absence of important features of joint attention. This overall finding suggests that joint attention may not be needed for successful word learning in two-year-olds and, as such, represents a noteworthy departure from existing theory and literature.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

- Akhtar, N. (2005a). Is joint attention necessary for early word learning? In C. Tamis-LeMonda, & B. D. Homer (Eds.), *The development of social cognition and communication* (pp. 165–179). Mahwah, NJ: Lawrence Erlbaum.
- Akhtar, N. (2005b). The robustness of learning through overhearing. *Developmental Science*, *8*, 199–209.
- Akhtar, N., Dunham, F., & Dunham, P. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language*, *18*, 41–49.
- Akhtar, N., & Gernsbacher, M. A. (2007). Joint attention and early vocabulary: A critical look. *Language & Linguistics Compass*, *1*, 195–207.
- Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, *72*, 416–430.
- Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*, 347–358.
- Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology*, *14*, 79–93.
- Akhtar, N., & Tomasello, M. (1998). Intersubjectivity in early language learning and use. In S. Braten (Ed.), *Inter-subjective communication and emotion in early ontogeny: Between nature, nurture, and culture* (pp. 316–335). New York: Cambridge University Press.
- Bakeman, R., & Adamson, L. (1984). Coordinating attention to people and object in mother–infant and peer–infant interactions. *Child Development*, *55*, 1278–1289.
- Baldwin, D. A. (1991). Infants' contribution to the achievement of joint attention. *Child Development*, *63*, 875–890.
- Baldwin, D. A. (1993a). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*, 832–834.
- Baldwin, D. A. (1993b). Infants' ability to consult the speaker for clues to word meaning. *Journal of Child Language*, *20*, 395–418.
- Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore, & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum.
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, R. N., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word–object relations. *Child Development*, *67*, 3135–3153.
- Baldwin, D. A., & Moses, L. M. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development*, *10*, 309–329.
- Baldwin, D. A., & Tomasello, M. (1998). Word learning: A window on early pragmatic understanding. In E. Clark (Ed.), *The proceedings of the twenty-ninth annual child language research forum* (pp. 3–23). Chicago, IL: Center for the Study of Language and Information.
- Behrend, D. A. (1990). Constraints and development: A reply to Nelson, 1988. *Cognitive Development*, *5*, 313–330.

- Behrend, D. A., Scofield, J., & Kleinknecht, E. E. (2001). Beyond fast mapping: Young children's extensions of novel words and novel facts. *Development Psychology, 37*, 698–705.
- Bruner, J. S. (1999). The intentionality of referring. In P. D. Zelazo, & J. W. Astington (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 63–75). Mahwah, NJ: Lawrence Erlbaum.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17–29.
- 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, Serial No. 255).
- Dunham, P. J., Dunham, F., & Curwin, A. (1993). Joint-attentional states and lexical acquisition at 18 months. *Developmental Psychology, 29*, 827–831.
- Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy, 9*, 327–339.
- Franco, F., & Gagliano, A. (2001). Toddlers pointing when joint attention is obstructed. *First Language, 21*, 289–321.
- Jaswal, V. K., & Markman, E. M. (2003). The relative strengths of indirect and direct word learning. *Developmental Psychology, 39*, 745–760.
- Kremer, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology, 10*, 41–63.
- Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*, 57–78.
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*, 121–157.
- Markus, J., Mundy, P., Morales, M., Delgado, C. E. F., & Yale, M. (2000). Individual differences in infant skills as predictors of child-caregiver joint attention and language. *Social Development, 9*, 302–315.
- Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3–4, Serial No. 220).
- Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and saliency cues. *Developmental Psychology, 35*, 60–68.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Rice, M. L., Huston, A. C., Truglio, R., & Wright, J. C. (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 from viewing. *Child Development, 59*, 420–429.
- Sabbagh, M. A., & Baldwin, D. A. (2005). Understanding the role of communicative intentions in word learning. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology* (pp. 165–184). New York: Clarendon Press/Oxford Press.
- Saxon, T. F. (1997). A longitudinal study of early mother-infant interaction and later language competence. *First Language, 17*, 271–281.
- Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69*, 309–320.

- Scofield, J., Williams, A., & Behrend, D. A. (2007). Word learning in the absence of a speaker. *First Language, 27*, 297–311.
- Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *Journal of Genetic Psychology, 164*, 54–71.
- Smith, C. B., Adamson, L. B., & Bakeman, R. (1988). Interactional predictors of early language. *First Language, 8*, 143–156.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore, & P. Dunham (Eds.), *Joint attention: Its origins and role in development*. Hillsdale, NJ: Lawrence Erlbaum.
- Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Tomasello, & E. Bates (Eds.), *Language development: The essential readings* (pp. 111–128). Malden, MA: Blackwell.
- Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development, 57*, 1454–1463.
- Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19*, 311–333.
- Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language, 4*, 197–211.