Cognitive Development 43 (2017) 190–200

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Research paper

# Language learning is hands-on: Exploring links between infants' object manipulation and verbal input☆

Kelsey L. West[a,*], Jana M. Iverson[b]

[a] Department of Psychology, University of Pittsburgh, 3215 Sennott Square, 210 S. Bouquet St., Pittsburgh, PA 15260, United States
[b] Department of Psychology, University of Pittsburgh, 3415 Sennott Square, 210 S. Bouquet St., Pittsburgh, PA 15260, United States

ARTICLE INFO

*Keywords:*
Language development
Motor development
Object manipulation
Verbal input

ABSTRACT

Research demonstrates that object manipulation may provide infants' with optimal sensory input for successful word learning to occur (Yu & Smith, 2012). Thus, it may be important to understand the relation between infant object manipulation and caregiver labeling in a naturalistic environment. The present study examined 13 parent-infant dyads longitudinally at 10, 12, and 14 months in their homes. The frequency and context of ideal labeling moments—-wherein infants are holding and visually attending to an object when the label is presented—-were examined. Results revealed that when infants were holding objects, caregivers' verbal input was less frequent but contained a significantly greater proportion of labels relative to moments when infants were not holding objects. Additionally, ideal labeling moments were more likely to occur during particular infant actions than during passive holding alone. Findings have implications for understanding the role that manual motor behavior plays in infants' early language environment.

## 1. Introduction

The notion of embodiment—that infants' actions on the world around them enhance learning—has a long history in developmental psychology (e.g. Gibson, 1988; Jean, 1954) and has received a surge of attention in the past decade (e.g. Libertus & Needham, 2010; Rieser, Lockman, & Nelson, 2005; Smith, 2005). It is theorized that the sensory input infants receive from self-produced action engenders learning about the environment in ways that passively experienced input does not. In particular, infants' manual exploration of objects has been shown to promote development in a wide range of domains, from object perception (Needham, 2000; Soska, Adolph, & Johnson, 2010) to social interactions (Libertus & Needham, 2011; Sommerville, Woodward & Needham, 2005). Notably, recent research has demonstrated that infants' object manipulation may be an important component of early language acquisition (Pereira, Smith, & Yu, 2014; Yu & Smith, 2012). This worksuggests that successful word learning is most likely to occur during moments when an infant is holding an object such that it dominates the visual field, and a caregiver provides the corresponding label. Understanding how these ideal moments for word learning unfold in the naturalistic environment may be useful in understanding processes that underlie early word learning. The present study investigates how infants' object manipulation behavior relates to the verbal input—in particular object labels—that they receive from caregivers, and how this relation changes as infants' object manipulation becomes progressively more sophisticated.

☆ This research was supported by a grant from the March of Dimes Foundation to S. Goldin-Meadow. Special thanks to the members of the Infant Communication Lab. at the University of Pittsburgh, and to the families and infants who participated in the research. Portions of these data were presented at the 2015 Biennial Society for Research in Child Development Meeting.

* Corresponding author at: Department of Psychology, University of Pittsburgh, 3215 Sennott Square, 210 S. Bouquet St., Pittsburgh, PA, 15260, United States.
*E-mail address:* kelsey.west@pitt.edu (K.L. West).

http://dx.doi.org/10.1016/j.cogdev.2017.05.004
Received 9 January 2016; Received in revised form 10 March 2017; Accepted 17 May 2017

At around one year of age, infants begin to produce their first words, arguably one of the most dramatic achievements of toddlerhood (Hollich et al., 2000; Schafer & Plunkett, 1998; Woodward, Markman, & Fitzsimmons, 1994). Given that word learning occurs in a cluttered environment and with few obvious or direct cues, this transition into language has led many researchers to investigate how infants are able to successfully map a heard word to the intended referent and encode this connection sufficiently for future recall. Work by Yu and Smith (2012) suggests that infants' object manipulation may play a facilitative role in this process. They propose that through *embodied attention*—a phenomenon in which held objects dominate the infant's view due to short arms and stature—infants are able to isolate the object from the environment and encode its label without the burden of selecting a referent from a cluttered visual array. This perspective underscores the importance of infants' object manipulation in the process of learning labels.

To test this hypothesis, they conducted a study in which infants wore a mini head camera and played with several novel toys with a caregiver in a lab setting. Prior to the play session, caregivers were taught novel labels for each of the items. They were then asked to interact naturally with the infant; no instructions were given regarding labeling the objects. Following the play session, an experimenter tested the infant's knowledge of the novel labels in a forced-choice task. Infants' performance in this task was then related back to the head-camera footage of the play session in order to examine whether and how infants' views of the objects differed for learned vs. non-learned labels. Results revealed that successful learning of word-object pairs was most likely to occur during moments of embodied attention, that is, when the label was presented while the infant was holding the object and it dominated the infants' view. Thus, these "right label at the right time" moments may be critically important for learning novel words.[1]

The embodied perspective builds upon a large body of work reporting that joint attention (i.e. moments when an infant and a social partner are simultaneously attending to the same object) is an important component of early word learning. This research has found that individual differences in the frequency and quality of joint attention between dyads—in particular caregivers' ability to follow the infants' attention—predict language development in the toddler years (e.g. Baldwin, 1995; Tomasello & Farrar, 1986; Tomasello, 1988). This work is closely related to the embodied perspective previously discussed, as joint attention is a necessary, but not sufficient, component of these "right label at the right time" moments. The essential distinction that embodied attention makes is that the quality of joint attention may also vary along another dimension: the sensory-motor experience of the infant. Thus, moments of joint attention that occur while the labeled-object is simultaneously being held by the infant (these "right label at the right time" moments) may provide the *optimal* sensory-motor input for infants to encode novel labels.

Additional support for the hypothesis that infants' object manipulation plays a role in early language development comes from naturalistic studies of infants' early word learning. In a seminal study by Nelson (1973), caregivers kept diary records of the words their infants produced and reported on them at monthly intervals. The first 50 words produced by the infant—regardless of chronological age—were examined. A clear pattern was evident in these early words: the majority were labels for objects that infants were likely to manipulate (e.g. "cup"). Although household objects that could not be manipulated were often present in the infants' environment (e.g. "table") they were far less likely to appear in the infant's repertoire than items on which the infant was able to act. Nelson notes:

Frequency of personal experience, exposure to words, strength of need, or desire cannot apparently explain the selection of these words…It is apparent that children learn the names of things they can act on, whether they are toys, shoes, scissors, money, keys, blankets, or bottles…With very few exceptions all the words listed are terms applying to manipulable or moveable objects [p. 31].

Although these findings are descriptive and causal inferences cannot be made, they are nevertheless consistent with the hypothesis that the ability to hold an object may aid infants in learning its label.

An important implication of this framework is that, in practice, it relies heavily on the coordination of infants' object manipulation and caregiver labeling. Whereas other well-studied mechanisms for infants' word learning depend on the *infant* to extract relevant information from patterns and cues in the environment, this embodied perspective relies on the *dyad* to coordinate infants' actions and caregivers' labels. Thus, if word learning is facilitated by these "right label at the right time" moments, as some studies suggest (Lifter & Bloom, 1989; Pereira, Smith, & Yu, 2014; Smith & Yu, 2008), it is important to understand: (a) how these moments unfold in naturalistic interactions between infants and caregivers; and (b) whether caregivers' verbal input differs—either in frequency or content—during infants' object manipulation, as opposed to moments when infants are not handling objects.

Insights into these questions come from studies investigating parental responsiveness. This work has yielded evidence that caregivers coordinate their verbal input with infants' object manipulation, and that this verbal input frequently consists of labeling utterances (Tamis-LeMonda, Kuchirko & Tafuro, 2013; West & Rheingold, 1978). West and Rheingold (1978) tested whether a wide variety of infant behaviors—including object manipulation—would elicit caregivers' verbal input. They brought mothers and their 12 month-old infants into an experimental suite equipped with microphones and cameras for 15 min and observed their naturalistic interactions. Experimenters coded these sessions for a variety of infant behaviors, including object manipulation, as well as maternal verbal input. If maternal input began after the onset of the infant's behavior or within 2 s following the behavior, it was coded as a response to the behavior. Overall, results showed that 70% of maternal utterances were produced in response to infant behaviors,

[1] Learning language is a complex problem, and so it likely requires a complex solution. It is probable that there are many mechanisms that account for world learning in infancy and they are not necessarily mutually exclusive to each other. Many other influential factors have been investigated in word-learning, such as cross-situational statistical regularities (Akhtar & Montague, 1999; Plunkett, 1997; Smith & Yu, 2008; Smith, 1995; Smith, Jones, & Landau, 1996), unlearned assumptions about language (Markman 1990; Markman 1991; Merriman, Bowman, & MacWhinney, 1989; Waxman & Kosowski, 1990), and children's already acquired vocabularies (e.g. Bergelson & Swingley, 2012), among others. Although the embodied perspective presented here is unlikely to account for word-learning completely, it does dramatically reduce the problem of referential ambiguity—the quintessential obstacle of early word learning—which makes it a particularly compelling factor in infants' early language acquisition.

suggesting that in general mothers were sensitive and responsive to their infants' actions. Additionally, when maternal responses were broken down by infant behaviors, the highest percentage of verbal responses occurred following infant object manipulation. These findings provide some evidence that dyads coordinate infants' object manipulation and caregiver verbal input.

Further support comes from recent work that specifically investigated caregivers' responses to infants' object manipulation (Tamis-LeMonda, Kuchirko & Tafuro, 2013). In this study, caregivers and their 14 month-old infants were recorded during two toy play scenarios: reading books and making necklaces. Instances of infants' object manipulation were identified, as well as caregivers' object manipulation and verbal input. Sequential analyses revealed that caregiver behaviors in general (both object manipulation and language) were likely to occur in response to infant object manipulation. Importantly, caregivers' referential language (e.g. "That's a bead"), but *not* regulatory language ("Stop that!"), was highly likely to occur in response to infants' object manipulation. This suggests that the types of verbal input caregivers provide differs depending on whether the infant is holding an object.

Taken together, the studies described above provide important information about how caregiver-infant dyads coordinate object manipulation and verbal input in naturalistic interactions. However, several important questions remain unaddressed. One is whether the increased "referential" language is *congruent* (i.e. corresponding to) the object the infant is holding. In order for these "right label at the right time" moments to occur, caregivers' labels must be consistent with the specific object the infant is holding and attending to, as opposed to other competing referents in the environment. Conversely, it may be equally as important to examine the frequency of other labeling moments (i.e. labels that do *not* correspond to either the infants' object manipulation, gaze, or both). It stands to reason that in order for these "right label at the right time" moments to be an effective means of pairing a heard word with a held object, the labels provided must correspond to the held object more frequently than they do not. Thus, investigating what proportion of caregiver labels correspond to infants' object manipulation and gaze is important in assessing the functionality of this embodied perspective.

A second question is whether all types of infant object manipulation are equal in eliciting these "right label at the right time" moments. Around the same time that infants are beginning to break into language, they are simultaneously changing the ways in which they act on objects in their environment. Infants' early object manipulation consists of mainly sensorimotor actions, in which the purpose of the action is to explore the properties of the object (e.g. mouthing, shaking, or touching with the fingers; Belsky & Most, 1981). At around one year of age, infants begin to produce functional actions (i.e. actions that have a conventional association with the object, such as pretending to talk on a toy phone) and these actions continue to increase in frequency until around a year and a half (e.g. Belsky & Most, 1981). It is possible that these different types of actions could be accompanied by different types of verbal input from caregivers. Additionally, it is likely that this relation between infants' object manipulation and caregiver labeling may depend on the developmental level of the infant. To illustrate, if an infant is predominately engaging in sensorimotor actions, the caregiver may frequently label during these actions. However, as the infant begins to demonstrate functional actions, the caregiver may become more selective, increasing labeling during these moments, and decreasing labeling during more basic sensorimotor actions. Thus, this relation may be dynamic and change across developmental time.

### 1.1. The present study

In light of the above considerations, this research was designed to investigate how caregivers coordinate verbal input with their infants' object manipulation and gaze—creating "right label at the right time" moments—and how this may develop over time in relation to developmental changes in object manipulation. Thus, we seek to address the following three research questions:

1. *Do caregivers vary their verbal input during moments of infant object manipulation in comparison to moments when infants are not holding objects?* It is predicted that caregivers will provide more verbal input during moments of infant object manipulation relative to moments when infants are not holding anything, and that the verbal input they provide will contain more object labels.
2. *How frequently do "right label at the right time" labeling moments occur as opposed to other labeling moments?* Previous work has investigated how infant object manipulation may elicit labeling from caregivers. However, in order for these ideal moments for infant word learning to occur, the label must correspond both to the infants' object manipulation and gaze. The present study will examine how caregivers' labels relate specifically to infants' object holding and eye gaze. It is predicted that these "right label at the right time" moments will occur more frequently than other labeling moments.
3. *Are "right label at the right time" moments more likely to occur during some types of object manipulation than others?* The present study will investigate whether caregivers' labeling varies in relation to how the infant acts on the object. It is predicted that more complex, developmentally advanced actions will elicit more of these ideal labeling moments than other less-advanced actions. Further, this interaction may change over time as infants' object manipulation develops.

## 2. Methods

### 2.1. Participants

This study included 13 mother-infant dyads observed at three time points: when infants were 10, 12, and 14 months old. This age range was chosen because it marks a period when infants are first beginning to break into word learning (e.g. Carey, 1978). Participants were part of a larger longitudinal study aimed at investigating the development of gesture and language from 10 to 24 months. Dyads were recruited from two Midwestern cities through word of mouth and from public birth announcements (see Iverson & Goldin-Meadow, 2005, for a detailed description of the study). All infants (7 female) were from middle- to upper-middle

class monolingual English-speaking families. All primary caregivers were mothers and all had at least some college education, ranging from undergraduate to graduate/professional degree experience.

### 2.2. Procedure

At each time point, infants and mothers were filmed for approximately 18 min (M = 18.21 min, range = 13.92–24.37) in their homes during a naturalistic toy-play session. Dyads were given a large bag of 18 age-appropriate toys to play with during the session (e.g., teddy bear, toy cars, a brush). This ensured that caregiver labeling and infant object manipulation were not influenced by variation in the objects available in the home.[2] Caregivers were given no instructions regarding labeling, and were told only to "play as you normally would".

### 2.3. Coding

Coding was conducted for caregiver verbal input, infant object manipulation, and infant gaze. This was completed by the first author and five secondary coders who were naive to the objectives of the study. Coding was completed using Elan software (www.lat-mpi.eu/tools/elan; Lausberg & Sloetjes, 2009).

#### 2.3.1. Caregiver verbal input

All caregiver speech produced during the play sessions was transcribed by one primary and two secondary coders. Speech was coded in units of utterances, which were defined as any sequence of words preceded and followed by silence, a change in conversational turns, or a change in intonational pattern (Crystal, 1985; Devescovi & Pizzuto, 1995; Iverson, Capirci, Longobardi, & Caselli, 1999).

Next, utterances were identified as either containing a label—an object noun—or not containing a label. In order be considered a label, object nouns must have been concrete (e.g. "phone" would count, but "imagination" would not), and could not be pronouns (Sandhofer, Smith, & Luo, 2000).

#### 2.3.2. Infant object manipulation

Object manipulation was coded in two passes through the videos. In the first pass, each bout of infant object manipulation was coded. Object manipulation was operationalized as any time when one or both of the infant's hands was completely grasping an object; touching an object with the fingers did not count as object manipulation. Additionally, coders identified which object was being held. If more than one object was being held at a time, coders annotated both objects.

The second pass focused on coding different action types within the object manipulation bouts. Four action types were coded: functional, sensorimotor, social bids with objects, and passive. If during a single bout of object manipulation the infant performed multiple different actions, the duration of each action was coded separately (e.g., one 10 s bout of object manipulation could be coded as 5 s of passive holding and 5 s of sensorimotor object manipulation). Instances of *functional* object manipulation were coded if the infant performed an action conventionally associated with the particular object (e.g., pretending to drink out of a cup; making a toy car drive). *Sensorimotor* object manipulation was coded when an infant's action involved exploration of object properties (e.g., mouthing; turning in hand to see more views of the object). *Social bid with objects* were coded when the infant lifted the object up or outwards toward the social partner. In order to be coded as a social bid, the action must have been accompanied by an additional communicative behavior, such as eye gaze toward the caregiver, vocalizations, or a postural shift (Iverson, Capirci, Volterra, & Goldin-Meadow, 2008; Iverson & Goldin-Meadow, 2005; Parladé & Iverson, 2011). This category included both shows and gives to the caregiver. Finally, *passive* object manipulation was coded when the infant was holding an object, but was not attending to it or performing any action with it. In instances when an infant was holding two objects and performing different manipulation types (e.g. passively holding one object while performing a sensorimotor action with another), both action types were coded and the corresponding object was specified.

#### 2.3.3. Infant gaze

Infants' visual attention is an important component of "right label at the right time" moments. Thus, in order to identify these moments it was important to establish whether caregivers' labels corresponded to infants' eye gaze. To this end, three primary coders watched all previously identified instances of caregiver labeling and identified whether or not the infant was looking at the referent at the time it was presented.

#### 2.3.4. "Right label at the right time" moments

Taking together the infants' eye gaze and object manipulation coding, every label the caregiver provided was categorized as corresponding to: (a) the infants' *gaze only:* the infant was looking at, but not holding the referent; (b) infants' *hands only:* the infant

[2] The decision to use one consistent set of toys across visits ensured that variability in infants' actions over time was not due to differences in objects (i.e. different objects may afford different actions). A trade-off to this decision is that it is not possible to control for exposure across visits. However, there was no indication that infants' became bored or uninterested in objects across sessions; in fact, the amount of time they spent holding the objects increased across sessions, $F(2.22) = 8.001$, $p = 0.002$.

was holding but not looking at the referent; (c) *both* gaze and object manipulation (i.e. "right label at the right time" moments); or (d) *neither*.

### 2.4. Reliability

Inter-rater reliability was assessed by having two pairs of secondary coders independently code 20% of the videos. (One pair coded caregiver verbal input and the other coded infant object manipulation). For caregiver verbal input, mean percent agreement for identification of utterances was 88% (range = 74–98.6%). Cohen's kappa statistics were calculated for classifying utterances as either containing a label or not ($\kappa$ = 0.95; range = 0.88–1.00), and for classifying utterances containing labels as either congruent or incongruent labels ($\kappa$ = 0.92; range = 0.82–1.00).

For infant object manipulation, mean percent agreement for identification of object manipulation was 83.3% (range = 72.4–98.7%). An intraclass correlation was calculated for the durations of object manipulation (M = 0.87; range = 0.73–0.95). Mean percent agreements[3] for identification of each action type (functional, sensorimotor, passive, and social bids) were 83.1% (range = 70.1–100%), 82.6% (range = 82.4–100%), 88.3% (range = 74.2–100%), and 80.3% (range = 66.7–100%) respectively. Intraclass correlations were also calculated for the durations of each action type (M = 0.90, 0.88, 0.94, and 0.90 respectively).

For infant eye gaze coding, Cohen's kappa was calculated for categorizing whether or not infants' gaze corresponded to the labeled referent. Interrater reliability was high; $\kappa$ = 0.86; range = 0.72–1.0.

## 3. Results

The aim of the present study was to examine the relation between caregivers' verbal input and infants' object manipulation behavior. In particular, we were interested in examining the frequency and context of optimal word-learning moments, in which infants' hands, gaze, and parents' labeling are all coordinated. Following preliminary analyses describing changes in caregivers' verbal input and infant object manipulation across time, we will present analyses addressing three primary research questions.

Prior to conducting analyses, data were checked for outliers. Any data point that was greater than three standard deviations above or below the mean was identified as an outlier, and was substituted with the next highest or lowest value in the distribution (Tabachnick & Fidell, 1996). Additionally, any variable that was computed as a proportion was arcsine transformed. All data were analyzed using version 22 of IBM SPSS for Windows.

### 3.1. Preliminary analyses

Before testing whether caregiver verbal input related to infants' object manipulation, we first examined the development of these behaviors independently. The first set of analyses examined how caregivers alter their verbal input over time, both in frequency and in content. The second set examined developmental changes in infants' object manipulation behavior. Descriptive statistics for caregiver verbal input and infant object manipulation variables are presented in Table 1.

#### 3.1.1. Caregiver verbal input

The first question of interest was whether caregivers increased the numbers of utterances they provided to infants across time. To control for variations in the total session durations, we calculated the rate per minute of caregivers' utterances at the 10, 12, and 14 month sessions. The data in Table 1 show that the rate per minute of utterances appears to be relatively unchanged across time points. This was confirmed by a one-way repeated measures ANOVA, which revealed no significant change in the rate of caregivers' utterances, $F(2,22) = 0.311$, $p = 0.736$.

Additionally, we examined whether the content of caregiver utterances might contain more labels at later time points by calculating the proportion of caregiver utterances that contained a label (i.e., the number of utterances containing a label divided by the total number of utterances). As is shown in Table 1, the proportion of utterances that contained a label appeared relatively stable across sessions; this pattern was confirmed by a one-way repeated measures ANOVA, $F(2,22) = 1.519$, $p = 0.241$.

#### 3.1.2. Infant object manipulation

The aim of the next set of preliminary analyses was to examine changes in infants' object manipulation behavior across development. We began by examining the proportion of the total session that infants spent in each type of object manipulation (functional, sensorimotor, social bids with objects, and passive) at each of the three time points. It was expected that across development, infants would show an increased proportion of time spent in functional and social bids with objects—which are more developmentally advanced—and would show decreases in developmentally prior behaviors (sensorimotor and passive object manipulation). These data are presented in Fig. 1.

[3] Typically, bouts of infant object manipulation lasted for long durations, with many action types occurring within each bout (e.g. an infant may hold a toy for a minute, and within that minute switch back and forth between passively holding it, mouthing it, and performing a functional action on it). Thus, coding this variable did not involve categorizing each object manipulation bout as a particular type, but rather identifying many action types within the duration of the object manipulation bout. For this reason, mean percent agreement of identification of action types, and interclass correlations for the duration of each action type were utilized for reliability purposes, rather than Cohen's kappa.

**Table 1**
Descriptive statistics for preliminary analyses: caregivers' utterances and infant object manipulation.

| | *Month* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 10 | | | 12 | | | 14 | | |
| | M | Md | SD | M | Md | SD | M | Md | SD |
| Caregiver utterances: | | | | | | | | | |
| Rate of utterances per minute | 14.92 | 13.38 | 5.41 | 15.02 | 15.07 | 5.44 | 13.79 | 13.06 | 6.33 |
| Proportion of utterances containing a label[a] | 0.29 | 0.32 | 0.11 | 0.31 | 0.35 | 0.09 | 0.34 | 0.31 | 0.1 |
| Infant object manipulation: | | | | | | | | | |
| Proportion of object manipulation[b] | 0.64 | 0.63 | 0.16 | 0.71 | 0.73 | 0.13 | 0.71 | 0.70 | 0.15 |
| Proportion of manipulation | | | | | | | | | |
| Functional | 0.05 | 0.04 | 0.05 | 0.07 | 0.07 | 0.04 | 0.14 | 0.11 | 0.08 |
| Sensorimotor | 0.38 | 0.35 | 0.13 | 0.33 | 0.34 | 0.09 | 0.28 | 0.29 | 0.09 |
| Social bids with objects | 0.02 | 0.01 | 0.02 | 0.02 | 0.03 | 0.02 | 0.03 | 0.02 | 0.03 |
| Passive | 0.17 | 0.18 | 0.09 | 0.27 | 0.25 | 0.14 | 0.24 | 0.23 | 0.11 |

[a] The proportion of total caregiver utterances that contained a label.
[b] The proportion of total session spent manipulating objects.

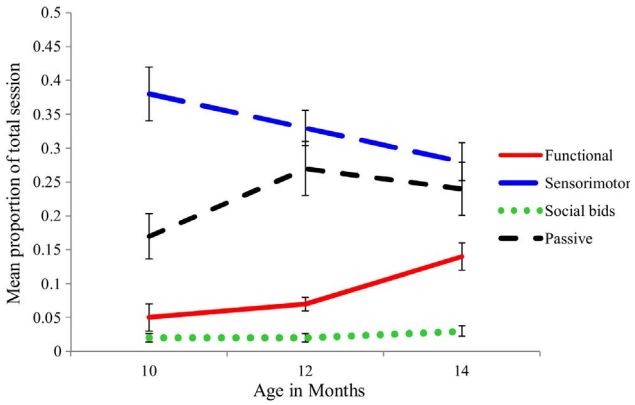


**Fig. 1.** Mean proportion of total session spent in each type of object manipulation.

The data in the figure indicate that infants spent the highest proportion of time in sensorimotor object manipulation, and this decreased across sessions as expected. Passive object manipulation was next most frequent, and showed a surprising pattern over time: it increased from 10 to 12 months, and then decreased from 12 to 14 months. Next, functional object manipulation showed the predicted increasing pattern over time. Finally, infants spent a very low proportion of time performing social bids with objects, and this remained stable over time.

A 4 (Action Type) by 3 (Age) repeated measures ANOVA revealed a significant main effect of Age, $F(2,22) = 8.001$, $p = 0.002$, with planned contrasts indicating that infants significantly increased the proportion of time spent manipulating objects overall—regardless of Action Type—from 10 to 12 months, $F(1,11) = 9.068$, $p = 0.012$; however, there was no significant change from 12 to 14 months; $F(1,11) = 0.033$, $p = 0.859$. Additionally, there was a significant main effect of Action Type, $F(3,33) = 69.524$, $p < 0.001$, which was qualified by a significant interaction between Age and Action Type; $F(6,66) = 5.015$, $p < 0.001$. Follow up simple effects analyses revealed that the increase in functional manipulation was significant from 10 to 12 months, $p = 0.02$, and again from 12 to 14 months, $p = 0.014$. Additionally, the observed decrease in sensorimotor manipulation was significant from 10 to 14 months, $p = 0.011$; however, this change was gradual and thus was not significant at the 12 month time point. The surprising pattern in passive object manipulation was also significant: passive object manipulation significantly increased from 10 to 12 months, $p = 0.005$; however, it did not change significantly at 14 months. No significant changes were found for social bids with objects across time.

In sum, these preliminary analyses revealed that caregivers' verbal input did not change significantly over time, either in frequency or in proportion utterances containing labels. However, infants did demonstrate developmental changes in object manipulation: infants spent more time manipulating objects overall, and their actions with objects became more advanced, with increases in functional and decreases in sensorimotor object manipulation. Below we examine whether and how caregivers' verbal input and infants' object manipulation may be temporally related.

**Table 2**
Descriptive statistics for caregiver utterances during infants' object manipulation and non-object manipulation: mean, median, and standard deviation.

| | *Month* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 10 | | | 12 | | | 14 | | |
| | M | Md | SD | M | Md | SD | M | Md | SD |
| Object Manipulation: | | | | | | | | | |
| Rate of utterances per min | 13.01 | 12.44 | 4.8 | 13.09 | 12.47 | 5.34 | 12.35 | 11.8 | 6.03 |
| Proportion of utterances containing a label[a] | 0.31 | 0.30 | 0.11 | 0.33 | 0.33 | 0.08 | 0.36 | 0.36 | 0.10 |
| Non Object Manipulation: | | | | | | | | | |
| Rate of utterances per min | 19.77 | 17.08 | 9.41 | 19.84 | 19.57 | 7.16 | 20.24 | 17.36 | 8.19 |
| Proportion of utterances containing a label[b] | 0.28 | 0.31 | 0.13 | 0.27 | 0.22 | 0.12 | 0.29 | 0.26 | 0.12 |

[a] The proportion of total caregiver utterances during infants' object manipulation that contained a label.
[b] The proportion of total caregiver utterances during infants' non object manipulation that contained a label.

### 3.2. Caregivers' verbal input during infants' object manipulation vs. non-object manipulation

A central aim of this study was to compare caregivers' verbal input during moments when infants were vs. were not manipulating objects. Our hypothesis was that caregivers would provide more verbal input during object manipulation, and that the input would contain a higher proportion of object labels.

The first analysis examined whether the rates of utterances that caregivers provided differed during infants' object manipulation and non-object manipulation. To test this, two variables were calculated: (1) the total number of caregiver utterances that occurred during object manipulation divided by the total time infants' spent manipulating objects; and (2) the total number of utterances that occurred during non-object manipulation divided by the total time infants' were *not* manipulating objects (i.e. the denominators for these two variables were different). These rates per minute were used in order to control for differences in the amount of time that infants spent manipulating objects. Descriptive statistics for each session are shown in Table 2; however, because means did not differ significantly across time points for both object manipulation, $F(2,22) = 0.208$, $p = 0.814$, and non-object manipulation, $F(2,22) = 0.051$, $p = 0.95$, data were collapsed across time points for the main analysis. It was revealed that caregivers provided a *higher* rate of utterances during infants' non-object manipulation ($M = 20.26$, $sd = 6.05$; range = 11.80–31.81) than during object manipulation ($M = 12.8$, $sd = 4.78$; range = 6.86–21.80). This pattern was confirmed with a paired-samples *t*-test; $t(12) = -5.005$, $p < 0.001$.

Although caregivers provided more verbal input overall during non-object manipulation than during object manipulation, it is possible that the input differed in *content*. In particular, we were interested in caregivers' production of object labels. To this end, we calculated the proportion of caregivers' utterances that contain labels separately for infants' object manipulation and for non-object manipulation (i.e. for object manipulation, this was calculated as: the number of labeling-utterances that occurred during infants' object manipulation divided by the total number of utterances that occurred during object manipulation). This allowed us to examine whether caregivers provided a higher proportion of labels during infants' object manipulation than during non-object manipulation. As in the previous analysis, means did not differ across time points for object manipulation, $F(2,22) = 2.224$, $p = 0.132$, or non-object manipulation $F(2,22) = 0.149$, $p = 0.863$, so data were collapsed across sessions. In line with our prediction, caregivers' utterances contained a higher proportion of labels during infants' object manipulation ($M = 0.33$, $sd = 0.6$; range = 0.23–0.43) than when infants were not holding anything ($M = 0.27$, $sd = 0.09$; range = 0.13–0.39). This pattern was confirmed with a paired samples *t*-test, $t(12) = 5.137$, $p < 0.001$. Taken together, these findings demonstrate that although caregivers are providing *more* verbal input during infants' non-object manipulation, the verbal input produced during object manipulation contains proportionately more labels.

### 3.3. Examining "right label at the right time" moments

Previous work has suggested that there may be ideal moments for infants to learn new words successfully: when infants are looking at a held object and a caregiver provides the corresponding label (Yu & Smith, 2012). The next set of analyses specifically examined these "right label at the right time moments" during infants' object manipulation, focusing on how they compare to other labeling moments.

The first question of interest was whether caregivers provided a higher proportion of ideal labeling moments—labels that correspond to the infants' object manipulation and gaze—during infant object manipulation, and whether this changed across development. To address this question, we calculated the proportion of total labels during infant object manipulation that corresponded to: (a) the infants' *gaze only* (the infant was looking at, but not holding the referent); (b) the infants' *hands only (*the infant was holding but not looking at the referent); (c) *both* (the infant was holding and looking at the referent, i.e. "right label at the right time" moments); or (d) to *neither*. These variables were calculated for each time point separately, and are presented in Fig. 2. Although data for labels that correspond to neither the held object nor gaze are represented in the figure, this category was excluded from analyses. This ensured that proportions across categories did not sum to 1.

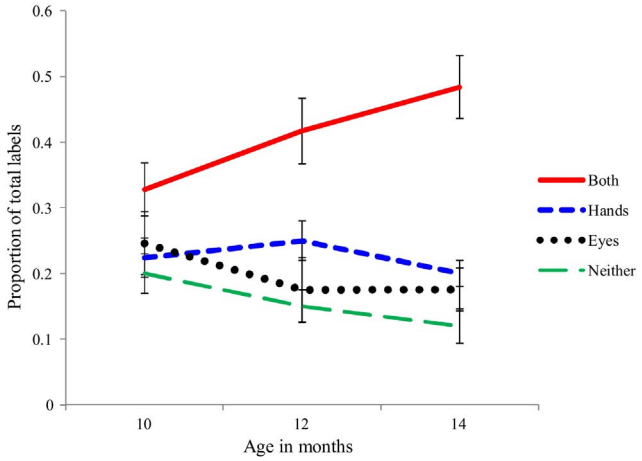


**Fig. 2.** Mean proportions of total caregiver labels that correspond to infants' hands only, gaze only, both hands and gaze, and neither during the entire session.

The figure suggests that during infant object manipulation, the majority of caregiver labels are coordinated both with infants' hands and gaze, and this appears to increase over time. A 3 (Age) x 3 (Congruence) repeated measures ANOVA was calculated, and there was a significant main effect of Congruence, $F(2,22) = 17.72$, $p < 0.001$. Planned pairwise comparisons confirmed that the proportion of labels that corresponded to both infants' gaze and hands was significantly greater than to infants' hands only, $p < 0.001$, and infants' gaze only, $p < 0.001$. Although there appears to an increasing trend in ideal labeling moments, the interaction between Age an Congruence did not achieve significance, $F(2,22) = 2.097$, $p = 0.097$.

### *3.4. The role of infants' actions with objects in these "right label at the right time" moments*

Next, we were interested in whether these "right label at the right time" moments were more likely to occur during some types of infants' object manipulation than others. Specifically, we predicted that caregivers would provide a higher proportion of ideal labeling moments during functional object manipulation—the most developmentally advanced action type—than during other types of object manipulation. For functional, sensorimotor, and passive object manipulation,[4] the proportion of total labels that corresponded to both the infants' hands and eyes was calculated (i.e. these "right label at the right time" moments). For example, for functional object manipulation this would be calculated as: [# of labels corresponding to infants' hands and gaze during functional/# of total labels occurring during functional]). These data are presented in Fig. 3.

To assess whether particular actions were more likely to elicit an ideal labeling moment than other actions, and whether this changed across development, a 3 (Action Type) x 3 (Age repeated measures ANOVA was calculated. As can be seen in the figure, across all action types, these ideal labeling moments appear to increase over time. This was confirmed by a significant main effect of Age, $F(2,22) = 57.44$, $p < 0.001$. Planned pairwise comparisons uncovered a significant increase in ideal labeling moments from 10–12 months, $p = 0.017$; however, the observed increase from 12 to 14 did not achieve significance, $p = 0.094$. Additionally, this analysis revealed a significant main effect of Action Type, with pairwise comparisons denoting that sensorimotor actions yielded the highest proportion of ideal labeling moments, significantly greater than both functional, $p = 0.005$ and passive object manipulation, $p < 0.001$. Functional actions had the second highest proportion of ideal labeling moments, which was significantly great than passive object manipulation, $p = 0.019$. These findings demonstrate that particular actions with object—most notably exploratory behaviors—are more likely to engender these "right label at the right time" moments than other actions.

## 4. Discussion

Laboratory-based research suggests that there are ideal moments for infants to learn novel words, and that these occur when the infant is holding and looking at an object and the corresponding label is presented simultaneously (Pereira, Smith, & Yu, 2014; Yu & Smith, 2012). The present study extends this work by characterizing how these moments unfold in a naturalistic setting. First, we examined whether bouts of infants' object manipulation were more likely to be accompanied by caregiver labeling than were moments of other infant activity. Additionally, we examined how these "right label at the right time" moments unfold in parent-infant interactions, how they relate to infants' actions, and whether these moments change across developmental time.

[4] Although coding included *social bids with objects*, this object manipulation type was excluded from analyses. Five infants did not demonstrate any social bids with objects at one or more time points, and so caregivers did not have an opportunity to provide verbal input for that action type. For this reason, only functional, sensorimotor, and passive object manipulation were included in the present analysis. Together, these types of object manipulation accounted for 96% of all time manipulating objects, and there was no significant change across time points; $\chi 2(2) = 2.167$, $p = 0.338$. Thus, the analyses presented here include the vast majority of time spent manipulating objects.

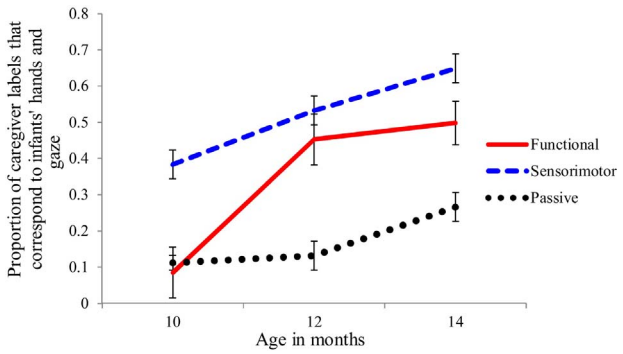


**Fig. 3.** Mean proportion of caregiver labels that correspond to infants' hands and gaze during each action type.

### *4.1. Do caregivers vary their verbal input during infants' object manipulation compared to moments when infants are not holding objects?*

Across the period of observation, caregivers provided lower rates of utterances during infants' object manipulation. However, this verbal input contained a significantly higher proportion of labels than did other moments. Thus, although caregivers' utterances were less frequent during infant object manipulation, their *content* may be more effective for early word learning, as evidence suggests that condensed and redundant verbal input (i.e. more repetitive labeling, fewer lengthy utterances) may facilitate early word learning (Murray, Johnson & Peters, 1990). Additionally, it is noteworthy that although the mean difference in the proportion of labels was significant, it was not very large: verbal input during object manipulation contained on average only about 5% more labels than during other moments. Although this is a modest difference, it may be meaningful as it unfolds over developmental time. As infants are consistently presented with more labels—and consequently more opportunities to learn object names—during object exploration, these moments may set the context for infants' word learning. Therefore, it is important to understand the dynamics of caregiver labeling during infants' interactions with objects.

### *4.2. Do labels presented during object manipulation correspond to the infants' held object and gaze?*

Another objective of this study was to characterize labels that occurred during infant object manipulation. Specifically, we were interested in whether these labels corresponded to the infants' held object and eye gaze (i.e. creating these "right label at the right time" moments) and whether the frequency of these moments changed over time. Findings showed that caregivers' labels were congruent to the infants' held object and gaze significantly more frequently than they did not. This suggests that during this time period—as infants are beginning to break into word learning (e.g. Carey, 1978)—caregivers are increasingly following in when infants' manual actions and visual attention are focused on an object. Thus, it seems probable that the coordination of infants' hands and eyes may provide redundant cues to caregivers that signify infants' engagement and a readiness to learn. This lends support for the feasibility of an embodied perspective. Because a great majority of labels presented during object manipulation correspond to both the held object and infants' visual attention, pairing a heard label with a held object will most likely result in successful referent selection. Given that the environment in which infants learn words is often noisy and cluttered (e.g. Quine, 1960), it is meaningful that held objects frequently co-occur with their labels. This perspective may offer a means of reducing the problem of referential ambiguity.

### *4.3. Do some infant actions engender more "right label at the right time" moments than others?*

Another possibility we considered was that the types of actions infants perform during object manipulation may also relate to the input they receive. In particular, as infants' object manipulation becomes more mature, the relation to caregiver labeling may also develop. Preliminary analyses revealed that infants did demonstrate clear changes in their actions with objects over time, increasing in functional actions and declining in sensorimotor actions. These patterns are consistent with previous research on the development of play (e.g. Belsky & Most, 1981). To explore whether these changes related to caregiver labeling, we tested whether some actions would yield higher proportions of these "right label at the right time" moments than others. Surprisingly, results revealed that sensorimotor actions yielded the highest proportion of ideal labeling moments than any other action type. Functional actions had the second-highest proportion of ideal labeling moments, which was significantly greater than moments when infants were passively holding objects. These differences between action types demonstrate that indeed caregivers are sensitive to infants' actions with objects, and coordinate labeling with specific infant actions.

The finding that sensorimotor actions yielded more ideal labeling moments than functional actions is surprising given that functional actions are regarded as more advanced. One possible explanation for this finding could be that as infants' actions become more elaborate, so does the verbal input they receive. Although the present study focused specifically on caregiver labeling, the verbal input infants' receive from caregivers contains other rich information about their environment. It may be that infants' functional actions elicit more verbal input about actions. For example, during one bout of functional play in which an infant held a toy phone to her ear, her mother said "Are you talking to daddy? Say hi daddy". Although these utterances do not contain the

corresponding label, 'phone', they do contain information about the function of a phone. Future studies should further investigate the relation between infants' actions and the types of verbal input they receive.

### 4.4. Limitations

There are several important strengths of the present study design, including micro-coding of both infant and caregiver behavior longitudinally within a naturalistic setting. However, it is important to discuss limitations of the sample included. The group of participants was small (N = 13) and demographically homogenous, including all white and highly educated mothers. Thus, it is important that future studies replicate these findings with more diverse samples. This may be of particular interest for infants from low socio-economic status (SES) households. In addition to a large body of research showing that caregivers from low-SES backgrounds provide less and less varied verbal input to infants (e.g. Leffel & Suskind, 2013; Hoff & Tian, 2005; Rowe, 2008), there is also recent work to suggest that infants from low-SES households interact differently with objects, including reduced overall manual exploration of objects, and developmental delays in the types of object exploration (Clearfield, Bailey, Jenne, Stanger, & Tacke, 2014). Given these findings, it is possible that the relation between infant object manipulation and caregiver verbal input differs compared to middle- to upper-SES peers. There is a need for studies examining this relation in low-SES samples, in order to delineate whether mechanisms underlying word learning may differ in this population.

### 4.5. Conclusions

These findings are consistent with an embodied perspective of early word-learning, suggesting that infants' actions influence their language environment. Within a naturalistic interaction in their own home, caregivers coordinated labeling with infants' manual actions. This provides ecological support for the feasibility of these "right label at the right time" moments as an underlying mechanism for early referent selection, and consequently language acquisition. Additionally, this study underscores the importance of examining the linguistic input infants receive from social partners within the context of their motor actions. Throughout the first few years of life, infants' developing motor abilities have a dramatic effect on the ways in which they interact with their environment. Given that there is ample evidence that caregivers' input is generally responsive to the infant (e.g. Tamis-LeMonda, Kuchirko & Tafuro, 2013), it is important to consider the transactional nature of these interactions. As infants begin to interact with objects in their surroundings in new ways, the feedback they receive may also develop. In other words, it may be overly simplistic to consider infants' language acquisition as the output of the verbal input they receive; rather, it may be better described as emerging and developing within a feedback loop in which the infant and caregiver are both contributing.

## Funding

This study was supported by a grant from the March of Dimes Foundation to S. Goldin-Meadow.

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*(57), 347–358.
Belsky, J., & Most, R. K. (1981). From exploration to play: A cross-sectional study of infant free play behavior. *Developmental Psychology*, *17*(5), 630.
Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3253–3258.
Carey, S. (1978). *The child as word learner*.
Clearfield, M. W., Bailey, L. S., Jenne, H. K., Stanger, S. B., & Tacke, N. (2014). Socioeconomic status affects oral and manual exploration across the first year. *Infant Mental Health Journal*, *35*(1), 63–69.
Crystal, D. (1985). Things to remember when transcribing speech. *Child Language Teaching and Therapy*, *2*(1), 235–239.
Devescovi, A., & Pizzuto, E. (1995). *Lo sviluppo grammaticale. G. Sabbadini (a cura di). Manuale di neuropsicologia dell'età evolutiva*, . Bologna: Zanichelli, 260–285.
Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, *39*(1), 1–42.
Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of communication Disorders*, *38*(4), 271–278.
Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., ... Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 1–135.
Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367–371.
Iverson, J. M., Capirci, O., Longobardi, E., & Cristina Caselli, M. (1999). Gesturing in mother child interactions. *Cognitive Development*, *14*(1), 57–75.
Iverson, J. M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication in Italian vs. American children. *First Language*, *28*(2), 164–181.
Jean, P. (1954). The construction of reality in the child.
Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*(3), 841–849.
Leffel, K., & Suskind, D. (2013, November). Parent-directed approaches to enrich the early language environments of children living in poverty. In Seminars in speech and language (Vol. 34, No. 04, pp. 267–278). Thieme Medical Publishers.
Libertus, K., & Needham, A. (2010). Teach to reach: The effects of active vs. passive reaching experiences on action and perception. *Vision Research*, *50*(24), 2750–2757.
Libertus, K., & Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. *Developmental Science*, *14*(6), 1355–1364.
Lifter, K., & Bloom, L. (1989). Object knowledge and the emergence of language. *Infant Behavior and Development*, *12*(4), 395–423.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77.
Markman, E. M. (1991). *Categorization and naming in children: Problems of induction*. Mit Press.
Merriman, W. E., Bowman, L. L., & MacWhinney, B. (1989). The mutual exclusivity biasn children's word learning. *Monographs of the Society for Research in Child Development*, I–129.
Murray, A. D., Johnson, J., & Peters, J. (1990). Fine-tuning of utterance length to preverbal infants: effects on later language development. *Journal of Child Language*, *17*, 511–525.
Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 1–135.
Needham, A. (2000). Improvements in object exploration skills may facilitate the development of object segregation in early infancy. *Journal of Cognition and Development*, *1*(2), 131–156.
Parladé, M. V., & Iverson, J. M. (2011). The interplay between language, gesture, and affect during communicative transition: A dynamic systems approach. *Developmental Psychology*, *47*(3), 820.
Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, *21*(1), 178–185.
Plunkett, K. (1997). Theories of early language acquisition. *Trends in Cognitive Sciences*, *1*(4), 146–153.
Quine, W. (1960). *Word and object*. Technology Press of Mass. Inst. of Technology.
Rieser, J. J., Lockman, J. J., & Nelson, C. A. (Vol. Eds.), (2005). *Action as an Organizer of Learning and Development. Vol. 33*Psychology Press.
Rowe, M. L. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*(01), 185–205.
Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language*, *27*(03), 561–585.
Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.
Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, *60*(2), 143–171.
Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In basic and applied perspectives on learning, cognition, and development. *The Minnesota Symposia on Child Psychology*, *28*, 1–32 Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*(1), B1–B11.
Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: motor skill acquisition facilitates three-dimensional object completion. *Developmental Psychology*, *46*(1), 129.
Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York, NY: HarperCollins College Publishers Inc.
Tamis-LeMonda, C. S., Kuchirko, Y., & Tafuro, L. (2013). From action to interaction: Infant object exploration and mothers' contingent responsiveness. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 202–209.
Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development*, *61*(5), 1461–1473.
West, M. J., & Rheingold, H. L. (1978). Infant stimulation of maternal instruction. *Infant Behavior and Development*, *1*, 205–215.
Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13 and 18-month-olds. *Developmental Psychology*, *30*(4), 553.
Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *25*(2), 244–262.