

## **UC Merced**

# **Proceedings of the Annual Meeting of the Cognitive Science Society**

### **Title**

Point-Light Displays Illuminate the Abstract Nature of Children's Motion Verb Representations

### **Permalink**

<https://escholarship.org/uc/item/1dp2p6wz>

### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

### **Authors**

Liu, Jing  
Golinkoff, Roberta M.  
Piper, Kim  
[et al.](#)

### **Publication Date**

2000

Peer reviewed

# Point-Light Displays Illuminate the Abstract Nature of Children's Motion Verb Representations

Jing Liu (jingliu@udel.edu)

Roberta M. Golinkoff (roberta@udel.edu)

Kim Piper (kimpiper@udel.edu)

School of Education, University of Delaware, Newark, DE 19716

He Len Chung (hlchung@astro.ocis.temple.edu)

Kathy Hirsh-Pasek (khirshpa@nimbus.ocis.temple.edu)

Christopher H. Ramey (rameychb@aol.com)

Department of Psychology, Temple University, 1801 North Broad Street, Philadelphia, PA 19122

Bennett I. Bertenthal (bbertent@dura.spc.uchicago.edu)

Department of Psychology, University of Chicago, 5801 South Ellis Ave., Chicago, IL 60637

## Abstract

The present studies explored children's representations of motion verbs first, in an elicited production study using point-light displays (lights attached to the joints of the human body) and then in a language comprehension task using the Intermodal Preferential Looking Paradigm (IPLP). Results indicated that children indeed ascribe meanings to the portrayals of actions in point-light displays. When children could not spontaneously produce the actual verb for an action, they used either a more specific or a more general familiar verb that was considered appropriate by adults. These findings suggest that even by the age of 3, children's representations of the actions that verbs label are amazingly abstract. This is the first set of studies to probe the nature of children's verb representations under circumstances where the portrayal of the action is stripped of an apparent agent, a location, instruments, or in some cases, the objects ordinarily required in transitive actions (e.g., a shovel in shoveling). Using point-light displays provides the field with a methodological tool for exploring the components of verb representation in both children and adults and for investigating children's verb acquisition.

## Introduction

The purpose of this paper is to take a closer look at verb representations and to examine the question of how 3-year-olds, who already know many verbs, extend familiar verbs to novel events depicted in point-light displays. To extend a verb to a point-light action, children must activate their action representations and map the patterns of light sequences to their verb representations. The use of point-light stimuli provides a stringent test of children's motion verb representations due to the removal of context.

Considering that the semantic structure of verbs provides a kind of conceptual frame for constructing larger linguistic units such as phrases and sentences, verb learning is of central importance for young language learners. Since motion words are among children's first words (e.g., Bloom, 1993; Smith & Sachs, 1990), two essential questions are

whether children can successfully form word-action mappings and how they extend an action verb.

Research has found that infants are keenly aware of movement and can use movement to individuate objects and actions (Sharon & Wynn, 1997; Spelke, Katz, Purcell, Ehrlich, 1994). There is also evidence that 18-month-olds can distinguish the causal actions of push and pull and map novel words to these actions after limited exposure to the word-action pairs (Casasola & Cohen, in press). However, it is not clear what perceptual cues children use to form word-action mappings and how they decontextualize the use of verbs to successfully extend verb labels to the actions they witness. For example, how do children understand that the word "jump" refers to a category of jumping motions that include different kinds of jumps made by the same actor (e.g., Elmo jumping off tables and chairs), and the same action performed by different actors (e.g., Elmo or Lala jumping off the chair)?

One line of research suggests that children use lexical principles to narrow down the possible meanings of words (Golinkoff, Mervis, & Hirsh-Pasek, 1994). According to Golinkoff, Mervis, Hirsh-Pasek, Frawley, and Parillo (1995), lexical principles guide children to learn not only object nouns but also action verbs. For example, the "principle of extendibility" posits that nouns not only label the original exemplar but a category of objects of "like kind." When transferred to the acquisition of action verbs, this principle states that an action verb, like a label learned for an object, can be extended to more than just the original action. Yet, what is the basis for verb extension?

Golinkoff et al. (1995) suggested that shape, or the overall motion configuration of the action, may provide a basis for children's verb extension. The shape of an action can be primarily affected by the path of motion (horizontal vs. vertical such as walking versus jumping), the involvement of arms and/or legs, and the type of instruments, if any, involved in the action. With respect to nouns, shape provides important information about the function and categorical membership of an object and serves as an

important basis for object noun extension (Landau, Smith, & Jones, 1998). It is quite likely that, when no other information is accessible, the shape of an action will provide defining information about the event, as “many verbs of motion have ... a typical appearance, a physiognomy” (Marconi, p. 159). This typical appearance of action may represent what Pinker (1989) labeled the “shape” of an event, and provide the basis of children’s verb extension. In fact, certain semantic factors such as MANNER, PATH, and CONVEYANCE, seem to be embedded in the “shape” of an event (Talmy, 1985). By attending to the overall shape or configuration of an action, children may ignore the context details of the individual event and construe the semantic invariant that a verb encodes. This would result in a more abstract and flexible representation of the action to which a particular verb can be mapped. For example, while the particular individual or object is not necessarily a part of the meaning of the verb FALL, the so-called shape of downward trajectory does represent the core, typical appearance of the action “falling.”

The claim that shape serves as an important basis for verb extension is in accord with Gibson’s (1966) view that in perceiving events we detect the “invariants” that persist from one event to another of the same type. The overall configuration of an action may be an invariant of the event. This claim is equally in line with Mandler (1992) who argues that events are stored as “image schemas” or “dynamic analog representations” abstracted from children’s interpretations of the spatial relations between objects. These meaningful image schemas help reduce the infinitely varying perceptual displays into a limited number of meaningful concepts that can be described by words.

Here we argue that children may use the “shape” of an action that loosely represents the invariant cues from one action to another as a basis for verb extension. It should be noted, however, that in using terms such as “invariant,” we are not assuming that the representations themselves are fixed and rigid. On the contrary, we posit that the shape of an action is a prototypical representation that is flexible enough to include actions that are similar to, but do not exactly match the original exemplar in terms of, say, the agent and location.

Point-light displays were first used to study adults’ event perception and biomechanical motion by attaching small lights to the head and main joints of an individual’s body and filming the person performing different actions against a dark background (Johansson, 1973). Because point-light displays are deprived of detailed contextual information such as the agent and location of an event, the information about an action given in such displays is expressed in the overall shapes of the light sequences. Previous studies using these moving light displays with infants and adults demonstrate the significance of prior knowledge in the perception of biomechanical point-light images (Bertenthal, 1993). For example, while 5-month-old infants can discriminate between a point-light walker shown in an

upright versus upside-down orientation, 3-month-olds do not demonstrate this sensitivity (Bertenthal, 1993). Thus, whether children can identify the actions depicted in point-light displays based on their previous experience provides a strong test of the hypothesis that toddlers use abstract, shape-based event representations to extend familiar motion verb labels. Research has found that 3-year-olds use many motion verbs, we therefore first investigated 3-year-olds’ ability to identify point-light depictions of human actions.

## Experiment 1

This study explored 3-year-olds’ ability to spontaneously produce a label for an action depicted in point-light displays. Since children do not encounter point-light displays in everyday life, to successfully complete this task, children must perceive the patterns of lights as meaningful action sequences, activate their verb representations for these actions, and ascribe meanings to these point-light sequences.

### Method

#### Participants

Thirty-eight children participated in the study. The final sample had 29 children, 16 boys and 13 girls, mean age = 3 years 7 months. Nine children were excluded from the data because 3 failed to produce a description for more than 3 of the actions and 6 either refused to talk to did not finish the study.

**Stimulus Videotapes** Biomechanical displays of motion verbs were created by videotaping a person in action with light-emitting diodes (LEDs) affixed to the major joints of the body (ankles, knees, hips, wrists, elbows, shoulders). These displays consisted only of a collection of white dots moving across a black screen. Figure 1 provides an example of a person walking, translated into frozen still images of point lights.

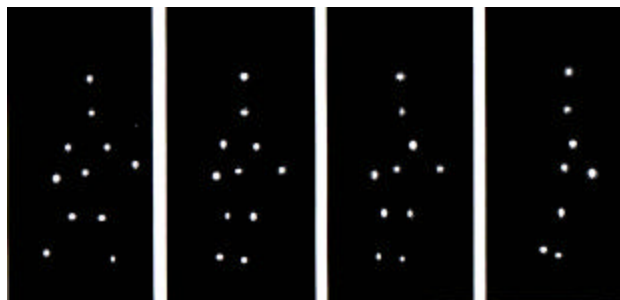


Figure 1: Canonical Point-light Walker

Each of the 8 actions was performed continuously for 3 seconds and repeated 4 times so that each action was

displayed for 12 seconds, followed by 5-second blank tape (Table 1). The actions were randomly ordered into two different sequences. To familiarize subjects with point-light displays, a point-light depicted cat walking from right to left on the screen was presented first.

### Procedure

The child sat on the parent's lap in front of a 32" TV. The parent was told not to say anything. Then the experimenter told the child they were about to see a fun videotape and would be asked to tell what they saw. After the cat display, the experimenter paused the VCR and prompted the child three times for a label for the action. If the child only produced an object label, such as "a cat" or "a doggie," the experimenter would probe again, "What was the cat/doggie doing?" If the child produced an unrelated answer such as "dots" or "snow," the experimenter would label the cat for the child and encourage the child to label the action. If the child did not produce an action label after being probed twice, the experimenter would produce the label, "Was the cat walking?" The same procedure was then repeated for each of the human actions except that the experimenter did NOT produce any description for the human action. If the child failed to produce a description for a human action after being probed twice, the experimenter went on to the next action. The experimenter made neutral comments as a response.

Table 1: Descriptions of the Actions in Point-light Displays

- \*Walking: Person walking.
- \*Dancing: Person twisting in place.
- Shoveling: Person bending at the waist and standing back up (without moving legs) as if shoveling snow.
- Picking flowers: Person bending at the waist and standing back up (without moving legs) as if picking up flowers from the ground and then putting them back in a handheld basket.
- \*Rolling: Person performing a somersault.
- \*Running: Person running.
- \*Skipping: Person skipping.
- \*Hopping: Person hopping on one foot.

\* Person moves diagonally from top left corner to bottom right corner across the screen.

**Data coding** Children's productions were divided into two groups: relevant vs. irrelevant. Words or phrases that did not indicate any motion were considered irrelevant, such as "have no feet." The rest were considered relevant. A group of college students (n = 15) were asked to rank the appropriateness of each relevant response on a 1-7 Likert scale after seeing each action in the point-light displays in the same way as children did.

Children produced a total of 110 different descriptions for the 8 actions. After eliminating 7 irrelevant answers and

combining 25 relevant answers by ignoring verb objects (i.e., "picking up snowballs" and "picking up stuff" were considered the same), there were 78 responses left. Each action had a range of 4-15 responses from children. These were listed on the ranking sheet. To examine the consistency of adults' ranking, 77 responses were used as foils for each other (the response "moving" was excluded as too general). The foils were distributed in ranking sheet for each action such that if adults were consistent in their ranking, these foils would receive on average lower rankings than the non-foils. The proportion of actual responses and foils for each action was 1:1, resulting in a total number of 154 responses for the 8 actions on the ranking sheet.

Table 2. Number of Children Who Gave Highly Appropriate Responses (standard deviation in parentheses)

Action	# of Children <sup>1</sup> /Total <sup>2</sup>	Mean Rating (>=5) <sup>3</sup>	Mean Rating (all responses)
rolling	10 / 25	6.33 (.70)	3.87 (2.00)
dancing	15 / 28	6.91 (.57)	3.17 (2.14)
picking flowers	10 / 25	6.41 (.68)	4.31 (2.22)
running	18 / 27	6.74 (.07)	4.65 (2.56)
walking	26 / 29	6.97 (.28)	4.77 (2.63)
hopping	16 / 27	6.78 (.70)	4.10 (2.42)
skipping	6 / 25	6.74 (.42)	3.46 (1.72)
shoveling	3 / 19	5.80 (.58)	2.69 (2.13)

1. Number of children who gave highly appropriate responses.
2. Total number of children who gave responses. Note that some children did not give responses for some actions.
3. Average rating for the highly appropriately responses.

### Results

Overall, most children produced labels for all actions after one prompt. Adults' rating for children's actual responses (M = 3.80, SD = 2.10) were significantly lower than for the foils (M = 1.45, SD = .84),  $F(1, 155) = 84, p < .001$ . Based on the adults' ratings, on average children's responses for most of the actions were considered appropriate (Table 2). Furthermore, 58% of children's responses were considered very appropriate (M = 6.73, SD = .58. All ratings greater than or equal to 5 were considered very appropriate). There was a significant correlation between the number of children who produced highly appropriate responses and adults' appropriateness rating of those responses,  $r = .36, p = .001$ , suggesting that the more appropriate a response was, the more children gave that response. However, few appropriate answers were given for the actions SKIPPING

and SHOVELING. Ten children were unable to produce a label for SHOVELING and 4 failed to produce a label for SKIPPING. The majority of others described SHOVELING as “dancing” or “exercising;” and SKIPPING as “running,” “jumping” and “walking.”

## Discussion

Findings of this study suggest that 3-year-olds can indeed perceive the abstract images depicted in point-light displays as meaningful actions based on the patterns of light sequences. Overall, 77% of the children produced responses for all actions. Given that these children had never been exposed to point-light displays, they must access abstract verb representations in order to identify these contextually bare actions. However, omission of objects in point-light displays might have increased the difficulty of identifying actions involving instruments, such as SHOVELING. In addition, some children's descriptions of SKIPPING overlapped with their description of WALKING and RUNNING, the overall shape of which are similar but for the differences in the movement of arms and legs. This interestingly indicates children's reliance on the overall configurations of these actions. Perhaps this problem, too, was caused by the nature of spontaneous production, a demanding task. Additional data of action labeling from 5 adults (100% correct) suggest that correct verbs can be produced for these abstract point-light displays. If children indeed possess abstract, shape-based verb representations and point-light displays can capture the properties of these actions, children should be able to map a familiar verb to these actions in a less demanding task.

## Experiment 2

The purpose of this study was to determine whether 3-year-old children could correctly extend familiar verbs to actions depicted in point-light displays in a comprehension task.

## Method

### Participants

Children's comprehension of at least 7 out of the 8 action verbs was confirmed with parents on phone calls. Out of the final sample of 32 subjects (44 were tested), 15 did not understand “skipping” and 3 did not understand “shoveling.” The age selected for testing was determined empirically from these phone. The final sample had 19 boys and 13 girls, age range from 2;11 to 3;2, mean age = 3;1.

**Stimulus videotapes** Two separate tapes were created, each containing half of the actions, paired in length and number of frames. The pairs of actions were created such that the lights appearing in the displays were balanced for size, number, brightness, and movement to ensure that each pair of actions was equally salient to children (Table 3). A female speaker recorded the linguistic stimuli in infant-

directed speech for all trials, as well as between the trials, on one track of the videotape.

### Procedure

Children were tested in the Intermodal Preferential Looking Paradigm (Hirsh-Pasek & Golinkoff, 1996). The child sat on the parent's lap in front of two 19" TV monitors. The videotapes were played in complete synchrony, accompanied by the linguistic stimuli which emanated from the center of the two monitors.

**Familiarization trial** The study began with a brief, 6-second trial during which a point-light display of a cat walking across the screen appeared simultaneously on both monitors for 6 seconds. The cat and its action (i.e., “See the cat walking!”) were labeled to give children some familiarity with interpreting the contents of these motion-specified stimuli.

**Salience trials** Two salience trials followed during which a pair of actions appeared simultaneously on both screens, one action on the left monitor and another on the right. Salience trials had three purposes: First, they showed children that contrasting events could appear on both screens at the same time. Second, they were used to calculate stimulus salience. Finally, they provided exposure to the names of the actions without telling children which screen either action was on, e.g., “Hey, one is walking and one is dancing!” Thus, children were directed to watch both screens.

**Test trials** Two test trials followed to see if the child could distinguish between the displays and successfully choose the action that matched the linguistic stimulus. Now the child was exhorted to watch the screen containing the labeled verb, e.g., “Look at dancing! See dancing?” The target action appeared on the same screen side for both test trials in a block, the same side as the two preceding salience trials.

**Intertrial intervals** Each trial was separated by a 3-second intertrial interval during which both screens went blank. A red light mounted centrally between the two televisions lit up during this time to attract children's attention to the center, off the screens. This practice ensured that children would not just remain on one screen for long periods of time, but would have to choose which screen to look at for each trial. The appropriate linguistic stimulus for the trial to follow was first heard during the intertrial interval, so that prior to each test trial, the child was directed to find the matching screen (e.g., “Can you find dancing?”).

**Apparatus and Data Coding** All equipment – except for the two 19" color monitors – was shielded from the child's view. The videotapes were shown on 3/4" video decks. A 1 KHz tone was recorded for the duration of each trial on the

second, inaudible channel of the videotape and was “read” by a specially designed tone decoder which functioned in two ways: 1) it turned the centrally-mounted red light on during the intertrial intervals, and off during the trials; and 2) it signalled the beginning and end of each trial to the computer (a PC computer).

**Dependent and Counterbalanced Variables** The dependent variable was the mean visual fixation time to the named action (the match) versus to the foil (the non-match) during each pair of the test trials. For each test trial, visual fixation time was collected starting during the intertrial interval from the point at which a child watched the center light for .3 seconds or more. Coding began during the intertrial interval because if a child failed to reach the .3 second intertrial interval criterion on a trial, that trial was not included in the data analysis (this occurred on only 5 trials, or 2% of the time). When a child missed a trial, his or her overall visual fixation mean to the match and non-match across the remaining test trials was substituted in that cell. Thus, each child contributed 8 data points to the analysis: the mean visual fixation time to the match and to the non-match for each of 4 pairs of test trials.

Four factors relating to order of stimulus presentation were counterbalanced across subjects: 1) the number of matches on a screen side; 2) the order of the matches; 3) the order of the two actions mentioned during the salience trial; and 4) the member of a verb pair labeled as the match.

## Results

Comparison of mean visual fixation times during the salience trials in the three-way mixed ANOVA (sex (2) X verb pair (4) X match versus non-match (2)) suggested that there were no a priori preferences for one action or another in any pair (Table 3). However, a significant difference was found between the mean visual fixation time to the match ( $M = 3.36, SD = 1.85$ ) vs. the nonmatch ( $M = 2.29, SD = 1.44$ ) during the test trials,  $F(1,3) = 48.81, p < .01$ . This effect was carried by the vast majority of children. Out of 32 subjects, 29 or 91% had mean visual fixation times in favor of the match for all of the four verb pairs.

## Discussion

When the motion-specified point-light images were presented in the IPLP (Hirsh-Pasek & Golinkoff, 1996), children who were 6 months younger than those in Experiment 1 demonstrated extension of all familiar verbs by watching the screen that matched the requested verb more than the nonmatch screen. Apparently, children could attend to the differences between the actions depicted in point-light displays in a comprehension task. Children could even map verbs to actions with which they were not familiar, such as SKIPPING. Perhaps their recognition of the unfamiliar verb was due to their successful mapping of a familiar verb to a familiar action presented as a comparison

in the IPLP. These results provided evidence that the combination of point-light displays and IPLP could be a powerful and sensitive tool to investigate children’s motion verb representations.

Table 3 Mean visual fixation times (in seconds) to the four stimulus pairs during salience and test trials (M = Match; NM = Non-match)

Verb pairs*	Saliency Trials	Test Trials**
	M (SD)	M (SD)
	NM (SD)	NM (SD)
Walking vs. Dancing	2.74 (1.46)	3.67 (2.29)
	2.82 (1.55)	2.22 (1.20)
Picking flowers vs. Shoveling	2.49 (1.15)	3.31 (1.69)
	2.84 (1.29)	2.37 (1.50)
Running vs. Rolling	2.39 (1.05)	3.40 (1.87)
	2.71 (1.18)	2.17 (1.37)
Skipping vs. Hopping	2.73 (1.49)	3.16 (1.56)
	2.53 (1.46)	2.05 (1.71)

\*The verb requested in each pair (i.e. match) was counterbalanced.

\*\* The match was watched more than the non-match for all test trials,  $p < .05$ .

## General Discussion

The purpose of these studies was to determine whether 3-year-olds could correctly extend familiar verbs to actions depicted in point-light displays. Experiment 1 found a majority of 3-year-olds could accurately identify the actions depicted in point-light displays. When children could not produce an accurate label for the point-light actions, they often used familiar verbs for actions that had similar overall shape to the target action. More than half of children’s responses (58%) were considered very appropriate by adults. Experiment 2 found that children who had just turned 3 could recognize all of the actions in the Intermodal Preferential Looking Paradigm (Hirsh-Pasek & Golinkoff, 1996), suggesting that the overall shape of an action could be a reliable basis for motion verb extension.

This is the first set of studies to suggest that young children can identify dynamic, complex events depicted in point-light displays and extend familiar verbs to the actions embedded in these events. Previous research focused on infants’ discrimination of familiar and unfamiliar biomechanical images (Bertenthal, 1993), and adults’ recognition of the familiarity and gender of a point-light walker (Kozlowski & Cutting, 1977). No work, however, had assessed the utility of these displays with young

children, let alone combining the stimuli with a language task. Point-light displays, which permit the use of dynamic events and contain so little contextual information, could serve as a critical tool to probe children's verb representations. Since the only available information in such abstract images was the "shape" of events, or an action's overall motion configuration, children's success in identifying these actions suggests that shape may be an important component of children's motion verb representations that guides motion verb extension.

We are not implying that children rely only on shape for motion verb extension. However, without disputing children's use of other complex verb learning processes (e.g., syntactic analysis), we underscore the flexibility of children's motion verb representations and the demonstration of their productive reliance on shape as one cue for the extension and categorization of motion verbs.

It is also important to note not all verb types can be illustrated through point-light displays. Verbs such as "see" or "think," for example, cannot be easily depicted through movie sequences, while motion verbs are ideal. Nonetheless, since motion verbs are generally among the first verbs acquired, future research employing these images may help uncover how early verb categories develop. Given that these contextually deprived displays can make contact with motion representations held by infants as young as 5 months, and that the children in this study could map a verb correctly to one of two choices, it appears that these stimuli could be used with older infants and children as well. Investigators could examine at what point young language learners are able to attach verb labels to these abstract images, and more specifically, when "shape" becomes a meaningful cue for extending novel verbs and forming motion verb categories. These types of research efforts can bring the critical study of verbs to the foreground, and help advance the understanding of fundamental verb learning processes that have for too long been neglected.

## References

- Bertenthal, B. I. (1993). Perception of biomechanical motions: Intrinsic image and knowledge-based constraints. In C. Granrud (Ed.), Carnegie symposium on cognition: Visual perception and cognition in infancy (pp. 175-214). Hillsdale, NJ: Earlbaum.
- Bloom, L. (1993). The transition from infancy to language: Acquiring the power of expression. New York: Cambridge University Press.
- Casasola, M. & Cohen, L. B. (in press). Infants' association of linguistic labels with causal actions. Developmental Psychology.
- Gibson, J.J. (1966). The senses considered as perceptual systems. Boston, MA: Houghton-Mifflin.
- Gibson, J.J. (1979). The ecological approach to visual perception. Boston, MA: Houghton-Mifflin.
- Golinkoff, R. M., Mervis, C. B., Hirsh-Pasek, K. (1994). Early object labels: The case for lexical principles. Journal of Child Language, 21, 125-155.
- Golinkoff, R. M., Hirsh-Pasek, K., Mervis, C. B., Frawley, W. B., & Parillo, M. (1995). Lexical principles can be extended to the acquisition of verbs. In M. Tomasello & W. E. Merriman (Eds.), Beyond the names for things: Young children's acquisition of verbs (pp. 185-221). Hillsdale, NJ: Lawrence Earlbaum Associates, Inc.
- Golinkoff, R. M., Jacquet, R., Hirsh-Pasek, K., Nandakumar, R. (1996). Lexical principles underlie verb learning. Child Development, 67, 3101-3110.
- Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The origins of grammar: Evidence from early language comprehension. Cambridge, MA: MIT Press.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. Perception and Psychophysics, 14, 201-211.
- Kowzowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. Perception and Psychophysics, 21, 575-580.
- Landau, B., Smith, L. B., & Jones, S. (1998). Object perception and object naming in early development. Trends in Cognitive Sciences, 2, 19-24.
- Mandler, J.M. (1992). The foundations of conceptual thought in infancy. Cognitive Development, 7, 273-285.
- Marconi, D. (1997). Lexical competence. Cambridge, MIT Press.
- Pinker, S. (1989). Learnability and cognition: The acquisition of argument structure. Cambridge, MA: MIT Press.
- Sharon, T. & Wynn, K. (1997). Individuation of actions from continuous motion. Psychological Science, 9, 357-362.
- Smith, C.A., & Sachs, J. (1990). Cognition and the verb lexicon in early lexical development. Applied Psycholinguistics, 11, 409-424.
- Soken, N.H. & Pick, A.D. (1992). Intermodal perception of happy and angry expressive behaviors by seven-month-old infants. Child Development, 63, 787-795.
- Spelke, E.S., Katz, G., Purcell, S.E., & Ehrlich, S.M. (1994). Early knowledge of object motion: Continuity and inertia. Cognition, 51, 131-176.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure and lexical forms. In T. Shopen (Ed.), Language typology and syntactic description: Vol. III. Grammatical categories and the lexicon. New York: Cambridge University Press.