The Effect of Semantic Representation on Toddlers’ Word Retrieval

Nina C. Capone
Karla K. McGregor
Northwestern University, Evanston, IL

Purpose: This study tested the hypothesis that depth of semantic representation influences toddlers’ word retrieval.

Method: Nineteen toddlers participated under 3 word learning conditions in this longitudinal study. Gestures cued attention to object shape (SHP) or function (FNC) in the experimental conditions. No semantic cue was provided under a control condition (CTL). Word learning conditions occurred on each of 3 days. On the 4th day, word retrieval was assessed across 3 levels of scaffolding (uncued picture naming, cued picture naming, picture recognition). Evidence of semantic representation was provided at fast and slow mapping intervals.

Results: Less scaffolding was necessary for word retrieval (uncued and cued naming) under experimental conditions than under the CTL condition. However, more SHP than FNC condition targets were retrieved for uncued picture naming. This latter difference may be related to the superior fast mapping of targets under the SHP condition. Toddlers stated object functions (slow mapping) comparably in the experimental conditions, but this was superior to CTL condition performance.

Conclusions: Word retrieval is a continuous behavior that is positively influenced by semantic representation. Semantic knowledge of objects can be enriched by shape or function gestures, thereby improving toddlers’ object word productions. Shape cues appear to be more effective for this purpose.

KEY WORDS: word learning, word retrieval, semantic representation, gesture, toddlers

The current study tested the hypothesis that depth of semantic representation is one factor that influences children's successful word retrieval (Bjorklund, 1987; Bjorklund & Schneider, 1996; Kail & Leonard, 1986; McGregor & Appel, 2002; McGregor, Friedman, Reilly, & Newman, 2002; McGregor, Newman, Reilly, & Capone, 2002; Plunkett, Karmiloff-Smith, Bates, Elman, & Johnson, 1997). Word retrieval failure, particularly when manifested as an error that relates semantically to the target word, has been linked to weak semantic knowledge of that word (weak semantic representation), whereas accurate naming is associated with a more elaborate knowledge base (rich semantic representation). However, studies supporting this relation have neither experimentally manipulated semantic enrichment or recall tasks nor fully controlled lexical factors known to influence retrieval such as phonological composition and frequency of exposure (e.g., Dell, 1990; Gershkoff-Stowe, 2002; Storkel, 2001). Furthermore, studies have examined preschoolers and school-age children. Semantic representation as a factor in toddlers’ word
retrieval, a group with new expertise in word learning, is less explored (but see Gershkoff-Stowe, 2001). The current study manipulated toddlers’ novel word learning to test the hypothesis that depth of semantic representation affects word retrieval.

**Semantic Representation and Word Retrieval**

Existing evidence that semantic representation shares a relation with word retrieval comes from children’s naming errors. Errors are most often logically related to targets and therefore reflect a speaker’s knowledge (Dell, 1990; Dell, Reed, Adams, & Meyer, 2000; McGregor, 1997). The relation between target and error can be phonological (e.g., _chicken_ to _kitchen_) or semantic (e.g., _key_ to _door_, _playpen_ to _crib_), although indeterminate (e.g., _thing_), visual misperception (e.g., _lollipop_ to _balloon_), and perseverative responses (i.e., the same word used to label two different objects within a defined time interval) can also occur. Semantic errors are prevalent and suggest a semantic component to retrieval failure when they occur. They may result from weak semantic knowledge or weak links between semantic knowledge and lexical labels (Gershkoff-Stowe, 2001; Lahey & Edwards, 1999; McGregor, 1997; McGregor, Newman, et al., 2002; Plunkett et al., 1997). For example, during the toddler’s word spurt, a time when weak entries are rapidly acquired, retrieval errors of known words increase (Gershkoff-Stowe, 2001; Gershkoff-Stowe & Smith, 1997). These errors are largely semantic and perseverative. As the child continues to amassed words, perseverative errors subside and semantic errors predominate (Gershkoff-Stowe, 2001). Connectionist simulations suggest that errors occur when connections between a lexical label and its semantic representation are not well established in memory (Plunkett et al., 1997).

McGregor and colleagues (McGregor & Appel, 2002; McGregor, Friedman, et al., 2002; McGregor, Newman, et al., 2002) studied the depth of semantic representation underlying naming responses in typically developing children and those with specific language impairment (SLI), a group that tends to have significant impairment (SLI), a group that tends to have significant impairment (SLI), a group that tends to have significant impairment (SLI). The current study manipulated provision of semantic cues while holding frequency of exposure and phonological composition of the lexical label constant. We then sought evidence of semantic representation at fast and slow mapping intervals. Fast mapping is the initial association of word and referent in memory (Carey, 1978). Fast mapped (and infrequently encountered) words are incompletely represented with limited semantic and lexical knowledge and few connections to other words in memory. That is, not much is known about a word after a single exposure. Although fast mapped words are weakly represented, this phase is important because it establishes a word-referent pair in memory. This leaves subsequent word-referent exposures available for enrichment. Semantic representations are gradually enriched during the slow mapping phase of word learning. This refers to the extended period of learning after a word is fast mapped (Carey, 1978). Slow mapping varies as a function of the
number and types of experiences individuals have with a given word. Experience is a key factor in creating a richer semantic network often superceding factors such as age and IQ (Bjorklund, 1987; Bjorklund & Schneider, 1996).

One consequence for the evolving nature of semantic representations is that more scaffolding may be needed to recall a word that is weakly represented. For example, in Dollaghan (1985), 2- to 5-year-olds identified the referent of a newly mapped word (/kub/) from an array of possible referents, but few recalled the lexical label to name it. A weak semantic representation may support word retrieval for a recognition task because cues that compensate for a weak representation are present (e.g., word, pictured referent, contrasting exemplars), but the probability of retrieving the word for naming (production) is lower (Bishop, 1997). Forced-choice recognition tasks test retrieval of the lexical label as it relates to its link with stored semantic information. Tasks assessing production (e.g., picture naming, free recall) provide fewer cues for retrieval, thus a representation needs to be rich for activation (Barsalou, 1999b; Plunkett et al., 1997). Production tasks moreover require retrieval and encoding of the lexical label (Gupta, 2004). In the current study, we used recognition and production tasks to tap both semantic representation and word retrieval.

Investigators have attempted to hasten the slow mapping process by making a word or referent more salient for the learner (e.g., Ellis Weismer & Hesketh, 1993; Gershkoff-Stowe, 2002; McGregor & Capone, 2004; Storkel, 2001). For example, Gershkoff-Stowe (2002) manipulated frequency of toddlers’ naming exposure and found that extra practice in naming resulted in fewer errors for a high practice item set. However, greater naming practice also provided more time to process picture stimuli, and the pictures themselves are a source of semantic information. Enriched semantic representations may have contributed to word retrieval performance. This could not be determined because semantic knowledge was not tested. In the current study, we both manipulated and assessed semantic knowledge during the fast and slow mapping phases of word learning to provide evidence of depth of semantic representation. Specifically, we asked participants to identify referents taught in semantically enriched and control conditions after one exposure (fast mapping interval) and to state semantic information about those referents after three exposures (slow mapping interval). We elaborate on this below.

**Gesture as a Cue to Word Learning**

Iconic gestures, also referred to as representational gestures, are manual or facial movements that carry meaning in their form and are believed to convey semantic information to communication partners (Acredolo & Goodwyn, 1996; Alibali, Bassok, Olseth Solomon, Syc, & Goldin-Meadow, 1999; Garber, Alibali, & Goldin-Meadow, 1998; Goldin-Meadow, 2000; McNeil, Alibali, & Evans, 2000; Morford & Goldin-Meadow, 1992). For example, holding two fingers up in a V can represent a rabbit, or extending an index finger up and making a circling motion can represent a fan (Acredolo & Goodwyn, 1996). Children attend to, interpret, and use gesture from infancy (for review, see Capone & McGregor, 2004). Kelly and Church (1998) found fourth graders recalled information expressed in peers’ gestures even though they were not instructed to pay attention to gesture. Toddlers studied by Morford and Goldin-Meadow and preschoolers studied by McNeil et al. were better able to follow novel spoken directions when supplemented with gesture. In the latter study, gesture was especially beneficial when directions were complex. This empirical evidence suggests that gesture scaffolds the child to higher performance levels when he or she is learning a new skill. This is true of word learning in particular. In Ellis Weismer and Hesketh (1993), iconic gestures enhanced children’s word learning over a no-gesture condition. For the at-risk infants studied by McGregor and Capone (2004), words modeled with gesture emerged in their spoken lexicon before other words.

In the current study, we used iconic gestures to cue semantic information during object label training. We compared two experimental conditions: one in which an iconic gesture provided a cue to object shape and another in which the gesture conveyed object function. Shape is a common basis of object label extension (i.e., determining whether a novel exemplar can be referred to by the same label as a known referent; Clark, 1973; Landau, Smith, & Jones, 1988), yet children also use object function to guide extensions when that knowledge is available to them (Booth & Waxman, 2002; Kemler Nelson, 1999; Kemler Nelson, Frankenfield, Morris, & Blair, 2000). In contrast, other object features, such as object size, do not influence extension decisions (e.g., Jones, Smith, & Landau, 1991). The relative value of shape versus function is debated. For example, Gershkoff-Stowe and Smith (2004) found parallel growth in toddlers’ shape bias and their productive noun vocabulary, yet infants studied by Booth and Waxman (2002) improved their novel noun extensions when provided function cues. Given existing evidence that both shape and function features are relevant to children’s mental representations of objects, both were included in the current study. We expected that both would enrich semantic representations over a no-gesture condition.

**The Current Study**

The current study tested the hypothesis that depth of semantic representation influences word retrieval.
Participants were toddlers, toddlerhood being a period when expertise in word learning has emerged and semantic naming errors are common. An experimental word learning paradigm manipulated the provision of semantic cues while controlling phonological factors, keeping word exposure constant, and counterbalancing stimuli across learning conditions. Semantic cues highlighted object shape or function and were provided by iconic gestures. The two cued conditions were created for comparison with each other and with a control condition. The latter served as a baseline measure of word learning. We sought evidence of semantic representation at both fast and slow mapping intervals by forced-choice object recognition and accuracy in stating objects’ functions, respectively.

Word retrieval was measured after a period of slow mapping. Traditionally, word retrieval is measured in a binary manner, retrieved or not retrieved. We reasoned that if semantic representations are graded, with weaker knowledge represented for some words and progressively richer representations for others, and if word retrieval is influenced by semantic representation, then word retrieval ability should also be graded. Specifically, richer representations should require less scaffolding for word retrieval, whereas a weak representation should require more scaffolding. Accordingly, our word retrieval task had three levels of scaffolding operationally defined by varying task support (uncued picture naming, cued picture naming, picture recognition). For each target word, we only analyzed the accurate response that required the least amount of scaffolding.

Our experimental questions were as follows: (a) Will a semantic representation be richer when a gestured semantic cue is provided during learning? (b) Will word retrieval require less scaffolding when learning involves a gestured semantic cue? and (c) Will children demonstrate parallel performance in semantic knowledge and amount of scaffolding required for word retrieval? We hypothesized that a semantic cue would enrich the semantic representation and positively influence word retrieval. We predicted a richer representation (relative to the control condition targets) at fast and slow mapping intervals characterized by (a) more words fast mapped and (b) more object functions stated under conditions of semantic enrichment. We predicted more words would be retrieved with less scaffolding when semantic cues were provided during learning than when no semantic cue was provided.

Method

Participants

Nineteen toddlers ($M = 28.70$, $SD = 0.99$, range = 27–30 months) were recruited from databases made available by the Child Language Laboratory and the Project on Child Development, both at Northwestern University. Participants were monolingual English-speaking toddlers (6 boys and 13 girls) from Chicago and its North Shore. They were of African American (5%), Asian American (11%), and Caucasian American (84%) backgrounds. The participants had no history of hearing impairment or developmental delays. Expressive vocabularies were typical ($M = 59.4$th percentile, $SD = 27.8$, range = 10th–99th) according to the MacArthur Communicative Development Inventory: Words and Sentences Form (MCDI; Fenson et al., 1993). Mothers were highly educated ($M = 17.60$ years, $SD = 1.35$), which is consistent with the economic status of the recruitment area. Exclusionary criteria were (a) MCDI score below the 10th percentile, (b) reported delays in development, or (c) familiarity with our objects. Former criteria were meant to ensure typical development; the latter to avoid interference in learning the nonce label (principle of mutual exclusivity; Markman, 1989). Initially, 23 participants were recruited, but 4 were excluded due to refusal to participate or speech-language delays. Number of excluded participants is consistent with other studies (e.g., Gershkoff-Stowe, 2002).

Stimuli

Stimuli were six objects from kitchen supply stores or thrift shops chosen because they were (a) novel to participants and (b) distinct in shape and function from each other. Three stimulus sets were created with two objects per set. A nonce label, nonce function, shape gesture, and function gesture were created for each object (see Table 1). These remained paired with the object throughout the study. Shape gestures were static iconic symbols, and function gestures were dynamic iconic symbols. Stimulus sets were balanced for gesture complexity (i.e., whether one or two hands were involved in the gesture) and the labels’ phonological characteristics (phoneme class, place of articulation, only CVC syllables). Relations between the target objects and other agents and entities in the environment were kept constant across objects to avoid a salient pairing that could differentially enrich semantic representations. Specifically, the first author conducted all sessions, and all objects’ functions were enacted with Play-Doh on the same toddler table.

Gesture Validation

To determine whether the iconic gestures accurately depicted objects’ shapes and functions, 10 adults from the Northwestern University community participated in a gesture-to-object matching task. Objects and their functions were demonstrated before adults manipulated them. A complete object array was then displayed along
with shape or function gestures presented in random order. The adults identified the object that matched each gesture. Mean accuracy in matching objects to the shape and function gestures was 95.7% (SD = 7.90) and 98.6% (SD = 3.80), respectively. A predetermined criterion of at least 80% accuracy for each gesture was reached.

**Word Learning Conditions**

Children learned words under three conditions: shape (SHP), function (FNC), and control (CTL). Spoken words labeled objects in all conditions, but shape gestures were paired with spoken words in the SHP condition and function gestures were paired with spoken words in the FNC condition. The order in which conditions were presented was counterbalanced across children but consistent across visits with any given child.

**Procedure**

The experimenter visited each participant at his or her home for 20 min on 4 separate days. One child had an equivalent schedule, but was seen at the university. Visits were videotaped for 18 participants; one mother preferred that her child not be videotaped.

The procedures were of three types. First, we were interested in manipulating type of semantic information while the child learned new words (word learning). Second, we probed semantic knowledge at fast and slow mapping intervals (object recognition probe, object function probe). Third, we examined the level of scaffolding required for word retrieval by manipulating task and provision of semantic cues (word retrieval probe). Each word learning condition (SHP, FNC, CTL) was presented at each of the first three visits. At Visit 1, an object recognition probe was administered after each condition to assess whether the word–referent pair was fast mapped. During Visits 2 and 3, an object function probe (Visit 2) and an object recognition probe (Visit 3) were administered between conditions, but these data are discussed elsewhere (Capone, 2003). At Visit 4, children were administered the word retrieval probe and lastly the object function probe. Preceding each probe task was a practice trial in which the experimenter referred to a familiar object (cup) with the word *cup* in combination with an abstract gesture for drink (thumb to lower lip) to illustrate that it was acceptable to use any form of communication. For all probes, noncontingent praise was provided for any response.

**Word learning.** Word learning was structured so that the experimenter first showed each object and demonstrated its function. Participants also manipulated each object to ensure familiarity with both shape and function. The experimenter then placed the object next

---

**Table 1. Novel objects, labels, functions, and gesture.**

<table>
<thead>
<tr>
<th>Label, function, and gesture</th>
<th>Apple divider</th>
<th>Jigger</th>
<th>Honey stick</th>
<th>J-bend pipe</th>
<th>Patty maker</th>
<th>Dumpling maker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonce label</td>
<td>/paxim/</td>
<td>/kas/</td>
<td>/n:aib/</td>
<td>/gef/</td>
<td>/dalan/</td>
<td>/w\lambda g/</td>
</tr>
<tr>
<td>Nonce function</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Stimulus set</td>
<td>Push down onto flattened Play-Doh leaving a distinct imprint</td>
<td>Pick up a ball of Play-Doh</td>
<td>Roll across flattened Play-Doh leaving a distinct imprint</td>
<td>Catapult Play-Doh from the shorter end when hitting the longer end</td>
<td>Flatten Play-Doh by turning the lid</td>
<td>Form a dumpling-shaped mass when closed</td>
</tr>
<tr>
<td>Nonce function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape gesture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Function gesture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The assignment of stimulus sets (1, 2, 3) to condition (shape, function, control) was counterbalanced across participants. Thus, not every participant saw each gesture.
to the child and stated, “It’s a [label]!” Labels were spoken nonce words (CTL) or spoken nonce words + gesture (SHP, FNC). For example, “It’s a /dæl/n/” or “It’s a /dæl/n/! + function gesture in the CTL and FNC conditions, respectively. The experimenter then labeled the object two additional times with spoken nonce words (CTL) or spoken nonce words + gesture (SHP, FNC; “You used the [label]!” and “Try my [label]”). The child manipulated it for another 60–90 s. Subsequently, “Have you seen my [label]?” was used to introduce objects. Objects were labeled three times during each learning session, summing to nine exposures per object by the study’s completion. Participants also heard nonce words during object recognition probes, but these were not paired with the object or gesture and exposure was comparable across conditions.

Object recognition probe. After each learning condition at Visit 1, children were presented a random array of four objects (two target objects, a cup, Play-Doh) and asked to identify referents when queried, “Where is the [spoken nonce word]?” To ensure that children remained on task and to minimize perseverative responding, participants identified the cup or Play-Doh at least once per condition. Percentage of objects identified was compared with chance levels of responding using the binomial distribution. Consistent with the principle of mutual exclusivity, chance was set at (.50) as random guessing would be between the two objects that were not already established in the lexicon. Post hoc analysis confirmed this. Erred choices were more likely to be the novel object rather than a familiar object (80% vs. 20% of trials, respectively).

Word retrieval probe. Word retrieval was assessed at Visit 4, which occurred, on average, 11.50 days (SD = 3.12) after the first visit, with a mode and median also at 11 days. Two participants were considered outliers, with fourth visits occurring at 5 and 21 days due to their availability. Outliers were not likely to have affected our interpretation of the relative benefit of one learning condition over another because participants served as their own controls. However, the outliers’ performance is analyzed first within the group and then separately to determine whether they fit with the group trends.

Participants were presented the word retrieval probe, which comprised three levels of scaffolding. Levels of scaffolding, proceeding from least scaffolded to most scaffolded, were uncued picture naming, cued picture naming (i.e., by iconic gesture by the experimenter), and picture recognition (see Table 2). Picture naming with uncued and then cued levels was probed first. Participants were given the opportunity to name pictures, and, if unable to retrieve the word, were then provided a semantic gesture cue. The most scaffolded level, picture recognition, was administered after all picture naming trials were completed.

Table 2. Word retrieval defined by level of scaffolding, retrieval response, and hypothesized semantic representation.

<table>
<thead>
<tr>
<th>Level of scaffolding</th>
<th>Retrieval response</th>
<th>Semantic representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncued picture naming</td>
<td>Spoken nonce label</td>
<td>Rich</td>
</tr>
<tr>
<td>Cued picture naming</td>
<td>Spoken nonce label when provided a semantic gesture cue</td>
<td>Moderate</td>
</tr>
<tr>
<td>Picture recognition</td>
<td>Pointing to pictured target when provided the spoken nonce label</td>
<td>Weak</td>
</tr>
</tbody>
</table>

Picture naming trials (uncued, cued) were embedded in a mailing game. Pictures were randomly presented for 5–10-s intervals, and the stimulus “What are we mailing?” elicited naming. If the child did not accurately label the picture, the experimenter provided a semantic cue (“It’s a [gesture]!”) by its shape (SHP condition) or function (FNC condition) gesture. For CTL condition objects, gestures were never paired with word or referent during learning, but one object’s shape and the other’s function gesture were provided to cue naming if toddlers failed to label the pictured object. Because the gestures were iconic, they served to scaffold responses even though they had not been directly taught.

Picture recognition trials were presented for all targets, regardless of uncued and cued picture naming performance. We did this for the experimenter’s ease of task presentation with this very young participant group and to ensure that the task’s experimental controls were maintained (e.g., target pictures occurred in all array positions to avoid the effect of perseverative responding). During these trials, children were presented an array of three pictures and queried, “Where is the [spoken label]?” Arrays were created from photographs of trained objects individually placed on poster board; the cup and Play-Doh were not pictured. Three pictures occurred on each poster board, and each object was pictured the same number of times across the poster boards. There were six poster boards so that each pictured object could be probed on a separate trial. Target position (i.e., for a particular trial) was counterbalanced across array positions to avoid perseverative responding. Trials were presented randomly until all targets were probed.

The word retrieval continuum was operationally defined by the amount of scaffolding needed for successful retrieval across these tasks. We categorized each
target word presented to the children \( (n = 6 \text{ per child}) \) by the accurate response that required the least amount of scaffolding; we then made between-conditions comparisons. This was a categorical analysis that compared each word retrieval response type as a proportion of all the trials presented in a condition. For example, if a child did not name a picture on first presentation (uncued picture naming), but named it when provided a gesture cue (cued picture naming) and identified it on the picture recognition trial, then word retrieval was categorized as cued picture naming only. Therefore, even though picture recognition trials were presented for all objects, these were only tallied in the analysis if the child accurately identified the pictured object and uncued or cued picture naming did not occur.

**Object function probe.** At Visit 4, after the word retrieval probe, knowledge of object function was probed. During pilot testing, toddlers had significant difficulty understanding queries about object shape and did not tolerate the time interval required to gather such data. Therefore, we probed only for object function. Objects were presented randomly after the word retrieval probe. Participants were queried “What do we do with this one?” while the experimenter showed each object. If the child responded, he or she was praised and the second object was presented. If the child did not respond, he or she was prompted, “Do we drink with this one? No, we drink with the cup ("cup" + gesture). What do we do with this one?” Spoken responses were transcribed verbatim. Accuracy was defined as spoken responses that were object specific. For example, functions such as “we scoop” for the /kas/ and “mix it” for the /doln/ were accurate because they were only characteristic of the target object’s function. Generic actions (e.g., “play”) were not specific to a particular object concept and therefore not tallied as accurate.

**Reliability and Fidelity of Treatment**

An independent coder who was blind to the study’s purpose and hypotheses recoded 37% of all dependent measures. Point-by-point agreement between independent coder and first author was 98.5% for the object recognition probe, 92% for the object function probe, 96% for the naming probe, and 100% for the comprehension probe. Interjudge reliability was also calculated for experimenter’s adherence to the protocol. An independent coder found the experimenter to be 100% accurate in application of experience with an object before providing a label for it and 98% accurate in providing type and number of labels (i.e., spoken, spoken + gesture, three labels).

**Dependent Variables**

Three dependent variables were measured. First, we compared percentage of accuracy in identifying a target object from an array of objects when asked for it by spoken label only (object recognition probe). Our second dependent measure compared level of scaffolding provided for an accurate retrieval response as a proportion of the total target words presented per condition (word retrieval probe). Therefore, for each target word, the accurate response that required the least amount of scaffolding was tallied (i.e., uncued picture naming, cued picture naming, or picture recognition). The third dependent measure was percentage of accuracy in stating object function (object function probe).

**Results**

**Object Recognition Probe: Fast Mapping**

After the initial object exposure, toddlers’ performance in identifying each object from an array of four objects was compared with that expected by chance (.50). An alpha level of .05 was used for all statistical tests. Only in the SHP condition did performance exceed chance levels with 68% of the word–referent pairs fast mapped, binomial \( (p = .01) \). The FNC and CTL conditions yielded at and below chance performance (42% and 34%, respectively). Shape cues paired with object labels facilitated fast mapping the word–referent pair.

**Word Retrieval Probe**

Participants’ word retrieval performance was tested across three levels of scaffolding, which proceeded from least to most scaffolded: uncued picture naming, cued picture naming, or picture recognition. The accurate response requiring the least amount of scaffolding was tallied in its respective scaffolding category. Consistent with Storkel (2001), participants had to produce at least two of the three target phonemes to be credited with an accurate uncued or cued naming response. Figure 1 illustrates the proportion of responses that were categorized at each scaffolding level between conditions. First note that the number of words retrieved, regardless of scaffolding, was comparable between conditions, with 84%, 91%, and 80% of the SHP, FNC, and CTL words retrieved, respectively, \( F(2, 15) = 1.00, p = .19 \). However, the amount of scaffolding needed for retrieval differed by condition. There were a total of 38 trials presented in each condition (19 participants \( \times \) 2 objects). Uncued responses made up 24%, 5%, and 3% of trials in the SHP, FNC, and CTL conditions, respectively. Cued responses made up 13%, 18%, and 3% of trials in the SHP, FNC, and CTL conditions, respectively. Picture recognition responses made up 47%, 68%, and 74% of trials in the SHP, FNC, and CTL conditions, respectively. Individual observation confirmed that the outliers’ performance was consistent with that of the group. The participant who was tested within 5 days demonstrated one uncued
picture naming response in the SHP condition; the remaining responses (n = 5) were categorized as recognition responses. The child who was tested within 21 days of the first word learning session demonstrated all recognition responses (n = 6).

Data were subject to a 3 (condition) × 3 (level of scaffolding) Friedman analysis of variance (ANOVA). Significant differences between response categories and learning conditions were detected, F(2, 16) = 3.12, p = .02, η² = 0.16. Post hoc testing by paired t tests with Bonferroni correction revealed performance in SHP (p = .02, one-tailed) and FNC (p = .01, one-tailed) conditions to be greater than that of the CTL condition, as predicted. There was no significant difference between SHP and FNC conditions (p = .65, two-tailed). In summary, toddlers stated more object functions in the experimental conditions than in the CTL condition.

### Discussion

The current study supported the hypothesis that depth of semantic representation influences word retrieval (Kail & Leonard, 1986; McGregor, Friedman, et al., 2002; McGregor, Newman, et al., 2002). As predicted, less scaffolding was necessary for word retrieval when words were learned under semantically enriched conditions. Words learned under CTL conditions were typically retrieved only within the most scaffolded task, picture recognition. Measures meant to provide evidence of semantic representation paralleled these findings. Toddlers knew more object functions in the semantically enriched conditions than in the CTL condition. Unexpectedly, word retrieval responses differed in SHP and FNC conditions, with more uncued naming in the former condition. This difference may be due to the fast mapping advantage participants gained under that condition. We explain this more fully below.

This work extends previous studies by (a) using an experimental word learning paradigm to test the causal relationship between semantic representation and success in word retrieval, (b) providing a graded analysis of word retrieval to illustrate the continuous nature of this behavior, and (c) providing evidence of semantic representation at both fast and slow mapping intervals. The results raise questions regarding the role of shape and function in facilitating object label learning and the role of fast and slow mapping in enriching semantic representation. These are addressed in turn.

### Semantic Representation Influences Word Retrieval

Previous research supports a positive relationship between depth of semantic representation and successful word retrieval for picture naming (McGregor, Friedman,
et al., 2002; McGregor, Newman, et al., 2002). Our experimental findings concur but also establish a direction of influence from semantic representation to retrieval of the word form. Holding lexical factors constant was a necessary component of the work because phonological composition and frequency of exposure also influence young children’s word retrieval (Gershkoff-Stowe, 2002; McGregor, Sheng, Grohne Reilly, & Keegan, 2004; Storkel, 2001). Furthermore, a word’s phonological composition influences semantic representation. When Storkel manipulated phonotactic probability (the likelihood of a given sequence of phonemes in the ambient language), preschoolers learned more commonly than rarely occurring CVC sequences. However, semantic representation was also affected. Semantic representations of common CVCs were considered “holistic” (Storkel, 2001, p. 1329), a conclusion supported by performance on several tasks. Whereas Storkel manipulated lexical factors and found effects on the semantic representation, we manipulated semantic factors and found an effect on lexical retrieval.

Studies have tested word learning with naming and recognition tasks (e.g., Dollaghan, 1985; Gray, 2004; Storkel, 2001), but retrieval responses have not been studied previously as a continuum operationally defined by decreasing task support (i.e., scaffolding). However, Dollaghan included a forced-choice label recognition task in addition to object naming and forced-choice object recognition tasks as measures of fast mapping. Her participants’ label recognition performance fell between that of object naming and object recognition. In other words, performance varied with amount of task scaffolding. This finding parallels our own: The graded nature of word learning can be tapped by tasks that vary in difficulty.

We hypothesized that depth of semantic representation positively influences word retrieval because richer semantic representations are distinct in knowledge and have stronger connection weights with the word. Evidence that more distinct semantic representations were established under the experimental conditions involves the superior performance in stating object function for items learned in the experimental rather than CTL conditions. Although we cannot directly measure activation weights, our word retrieval analysis suggests that words learned under SHP conditions reached activation thresholds because of the larger number of unscaffolded responses compared with other conditions; cued responses of the FNC condition suggested that thresholds were reached less often when compared with SHP condition targets but more often than the CTL condition targets. The weakest activation characterized more CTL condition targets, given the amount of scaffolding needed to retrieve those items. Future work that uses reaction time paradigms would reveal the time course of activation. Other sources of retrieval error were not studied. Failure to access a stored phonological form, visual misperception of pictured objects, and perseverative responding could also have been sources of word retrieval failure. However, we expect these errors to be comparable across conditions and therefore were not likely to influence our interpretation of the results.

In summary, the number of words established in memory, regardless of scaffolding, was comparable between conditions, but the amount of scaffolding necessary for word retrieval differentiated SHP from the FNC and CTL conditions and FNC from the CTL condition. It was not the case that at the end of the study more words were stored in memory under SHP and FNC conditions; rather, toddlers had richer knowledge of words when compared with the CTL condition. Richer semantic representations in the former conditions resulted in words requiring less scaffolding for retrieval. Failure to retrieve words for naming in the CTL condition was not due to missing representations but to weak representations.

**Shape Versus Function Cues and Object Label Learning**

An unexpected finding emerged from this study, namely, that shape and function cues led to different results in fast mapping but comparable slow mapping of object function. We attempt to understand these findings within an associationistic framework (Barsalou, 1999a; Smith, Jones, & Landau, 1996; Smith, Jones, Yoshida, & Colunga, 2003). According to Barsalou and colleagues, mental representations of objects are an integration of multimodal characteristics including lexical label, perceptual features (shape, function, related entities), and proprioceptive information extracted from direct object experience. From this theoretical account follow several possible explanations for the superiority of shape over function cues in facilitating fast mapping. Recall that in all conditions, participants’ observation and play allowed enactment of function prior to an object being labeled. The experimenter then provided the object, word label, and semantic cue (shape or function) simultaneously. We hypothesize that shape cues were more effective than function cues because they (a) capitalized on shape as a statistically reliable indicator of category (i.e., objects that share the same name tend to be similar in shape; Gershkoff-Stowe & Smith, 2004; Smith et al., 1996; Smith et al., 2003) and/or (b) were influenced by our methodology. These explanations are discussed in turn.

First, shape cues are a reliable indicator of object category. When objects are labeled, children use shape information to extend that label to a novel exemplar (Smith et al., 1996; Smith et al., 2003). The degree to which a toddler relies on shape-based extensions in laboratory experiments predicts object vocabulary
growth outside of the laboratory (Gershkoff-Stowe & Smith, 2004). Also, Jones (2003) found that late talkers, children defined by their small vocabulary size, do not demonstrate a shape bias. Kemler Nelson et al. (2000) found toddlers to make function-based extensions when given longer response intervals but to rely on shape information when time to respond was shorter, suggesting primacy of shape over function in children’s object extensions. Smith, Jones, Landau, Gershkoff-Stowe, and Samuelson (2002) strengthened the child’s shape bias for 17-month-olds by highlighting shape across same-named targets. They also contrasted those targets with objects that differed in shape and name. Children’s shape-based word extensions rose under these conditions, and there was parallel growth in object vocabulary outside of the laboratory when compared with children without shape bias training. We hypothesize that during fast mapping, our shape cues capitalized on the shape bias, strengthening its effect when paired with object and label.

Second, methodological decisions may have contributed to differences in the effectiveness of shape and function cues. These include participants’ age, enactment of all objects’ functions on a single entity (Play-Doh), and modality and timing of semantic cues. Shape and function vary in saliency across development, and our participants represented a tightly defined age group. Although toddlers can consider function in their word extensions, they tend to rely more heavily on shape at this time. With age, the shape bias weakens (e.g., Imai, Gentner, & Uchida, 1994). Therefore, shape cues may have been more effective than function cues because we happened to target an age at which shape cues are particularly accessible. Another potential methodological influence stems from our decision to demonstrate the function of each object on a common medium, Play-Doh. This may have rendered the object functions less distinct than their shapes. However, the participants’ ability to describe unique functions in both SHP and FNC conditions weakens this explanation.

Finally, the timing of gesture cue presentations during learning exposures may have benefited participants’ interpretation of shape cues. Shape cues were provided while viewing the object. Both being visual stimuli, there was a modality match for toddlers to interpret. Provision of function cues required participants to integrate the visual information presented and proprioceptive information stored during object enactment. Furthermore, there was a short delay between the uptake of proprioceptive information and provision of gestures. Function cues came after the toddler used objects, whereas shape cues were provided simultaneously while toddlers viewed objects. Even though function cues ultimately enriched semantic representations over the CTL condition, this may have taken more time than in the SHP condition.

It is also possible that shape cues were more salient because they were held longer than function cues. However, we have suggestive evidence to the contrary. We compared the participants’ imitations of shape and function gestures across word learning sessions. Although imitation of the gestures did not occur often, we found participants’ imitation of shape and function gestures to be comparable ($p = .26$). The data suggest that shape and function gestures were perceived equally well by the toddlers.

Though lacking a definitive explanation, the evidence is clear that provision of shape cues facilitated fast mapping more than provision of function or no semantic cues. Although weakly represented, a fast mapped word allows for subsequent exposures to be used for enrichment and is therefore a critical step in word learning. In Gray (2003), fast mapping accounted for “a significant but relatively small amount of variance” (p. 64) associated with preschoolers’ later comprehension and production of new words. Our data suggest that shape cues may boost toddlers’ fast mapping of object labels and their referents, hastening entrance into the slow mapping phase.

Children ultimately learned objects’ functions comparably in the SHP and FNC conditions despite only receiving cues to function in the latter condition. Chaigneau and Barsalou (in press) argued that object representations are a relational system between word, object parts, and object function. Across word extension studies, when children fail to appreciate the relation between function and object parts (shape), they also fail to use function as a basis for word extension; instead, they rely on shape. However, when the relation between object shape and function are transparent, children transcend physical features and consider function in their extensions. Whereas our shape cues may have improved association between word and object at the fast mapping interval, shape cues may have also helped toddlers appreciate the transparency between shape and function. In this way, shape cues may mediate the relation between function and word as well by facilitating creation of a holistic (i.e., relational) representation of the object. However, because we were unable to probe shape information, it is not known if function cues in the FNC condition influenced shape learning or whether this was a directional relation from shape cues to function learning. The current study was not designed to test this hypothesis; therefore, our data are only suggestive.

**Gesture Enhances Language Learning**

There is a growing body of evidence that gesture facilitates language development, including word learning (Capone & McGregor, 2004). Iconic gestures convey semantic information through “hand shape, placement
or motion” (McNeil et al., 2000, p. 132), so adults’ gestures may reinforce the semantic content of speech. When gesture co-occurs with speech in a child’s communication, the child’s own gestures may serve to externalize a visual representation or draw attention to important aspects of a problem they are discussing, thereby freeing neural resources for other cognitive processes (e.g., Alibali & DiRusso, 1999; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; McNeil et al., 2000). Our findings fit well with these previously articulated ideas. Specifically, we used iconic gesture to convey salient semantic information of novel spoken language (i.e., the word form) and found that children stored more semantic information than in the no-gesture condition. Perhaps our gestures drew attention to an important aspect of the word learning problem (shape, function, or both), thereby reinforcing salient semantic content of the spoken language.

**Clinical Implications**

Two clinical implications follow from this work. First, when assessing children’s word knowledge, the clinician is wise to recognize the graded nature of word knowledge. Words are not simply known or unknown but known to different degrees. Provision of a scaffold of increasing support for children’s performance on vocabulary tasks may reveal the degree of knowledge that a child has for any given vocabulary item.

Second, and more speculative given that the current study was limited only to children developing within normal limits, is an implication for intervention. Methods that facilitate semantic learning of object concepts hold promise as a means of improving word retrieval. For children with language impairments characterized by poor phonological memory (Gathercole & Baddeley, 1990), using gesture to enrich semantic knowledge may be especially prudent as gesture involves modalities—visual and motor—that may be relative strengths for these children. Because language impairments are often first identified on the basis of weak vocabulary skills and because, with development, these impairments transcend vocabulary skills to affect both grammar and reading (e.g., Rescorla, 2002), interventions that facilitate early word learning and usage are especially important.

**Conclusions**

In the context of previous studies, we conclude that the amount of scaffolding needed for word retrieval appears to be, in part, a function of the depth of semantic knowledge stored in memory. Word retrieval is not a binary behavior, learned for comprehension or production, but rather a more continuous one that is, in part, positively influenced by the graded nature of semantic representation. Our data suggest that knowledge can be enriched at fast and slow mapping intervals, and that the effect may be cumulative. Finally, provision of shape cues is an effective means of facilitating growth in toddlers’ object word productions.

**Acknowledgments**

This work was part of a doctoral dissertation completed by the first author and directed by the second author while both were at Northwestern University. We wish to thank Sandra Waxman, Susan Goldin-Meadow, and Steve Zecker for their contributions to project development and data analysis, as well as Julia Evans for additional discussion of our findings; Anne Graham, Kristy Grohne, Robyn Newman, Renee Reilly, and Li Sheng for discussion at every phase of the work; and Rosi Carr and Kanika So for careful reliability coding. We thank the families who participated in the study for welcoming us into their homes. The first author wishes to acknowledge support during data collection and analysis from a dissertation year fellowship awarded by Northwestern University. The second author was supported by Grant R29 DC 03698 from the National Institute on Deafness and Other Communication Disorders during the data collection and analysis of this project.

**References**


Received April 19, 2004
Revision received November 24, 2004
Accepted April 19, 2005
DOI: 10.1044/1092-4388(2005/102)

Contact author: Nina C. Capone, who is now at Department of Speech-Language Pathology, School of Graduate Medical Education, Seton Hall University, 400 South Orange Avenue, South Orange, NJ 07079.
E-mail: caponeni@shu.edu

Karla K. McGregor is now at the University of Iowa.