



Published in final edited form as:

J Child Lang. 2009 September ; 36(4): 807–828. doi:10.1017/S0305000908009173.

Gesture as a support for word learning: The case of *under**

KARLA K. MCGREGOR,

The University of Iowa

KATHARINA J. ROHLFING,

Bielefeld University

ALLISON BEAN, and

The University of Iowa

ELLEN MARSCHNER

The University of Iowa

Abstract

Forty children, aged 1;8–2;0, participated in one of three training conditions meant to enhance their comprehension of the spatial term *under*: the +Gesture group viewed a symbolic gesture for *under* during training; those in the +Photo group viewed a still photograph of objects in the *under* relationship; those in the Model Only group did not receive supplemental symbolic support. Children's knowledge of *under* was measured before, immediately after, and two to three days after training. A gesture advantage was revealed when the gains exhibited by the groups on untrained materials (but not trained materials) were compared at delayed post-test (but not immediate post-test). Gestured input promoted more robust knowledge of the meaning of *under*, knowledge that was less tied to contextual familiarity and more prone to consolidation. Gestured input likely reduced cognitive load while emphasizing both the location and the movement relevant to the meaning of *under*.

Gesture contributes to three mutually dependent processes: learning (Goldin-Meadow, 2000), thinking (Alibali, Kita & Young, 2000) and communicating (Kendon, 1994). Gesturing in concert with spoken language may reduce cognitive demands on these processes by allowing two different sorts of representational systems, visual and verbal, to share the load (Goldin-Meadow, 2000).

Consider first the effect of gesture on listeners. Adults' comprehension and retention of spoken messages are better when those messages are augmented by symbolic gestures that reinforce the meanings being expressed than when they are not. This is especially true under difficult conditions such as listening in the context of noise (see review in Kendon, 1994). Children too are better able to comprehend spoken messages when these are accompanied by symbolic gestures than when they are accompanied by conflicting gestures or no gestures (McNeil, Alibali & Evans, 2000). For example, preschoolers asked to follow directions stated via complex clause structures performed better when gestures reinforced the meaning

*This research was generously supported by grant NIH-NIDCD 2 R01 DC003698 awarded to the first author. We thank the children and parents who participated, Amanda Murphy, who assisted with scheduling, Jessica Werts, who assisted with data entry, and Nina Capone, who provided helpful comments on an earlier draft of this manuscript. Some of the data in this paper were presented at the 2006 and 2007 Symposium on Research in Childhood Language Disorders, Madison, WI, USA, by the first author and the third author, respectively.

© 2008 Cambridge University Press

Address for correspondence : Karla K. McGregor, PhD, 121c Speech and Hearing Center, Department of Speech Pathology and Audiology, University of Iowa, Iowa City, IA 52242. karla-mcgregor@uiowa.edu.

of locative words in the directions (e.g. the word *up* with a gesture to indicate ‘up’) than when the gestures conflicted with the meaning of the words (e.g. the word *up* with a gesture to indicate ‘down’). Importantly, children do not necessarily derive benefit from reinforcing gestures when the spoken message is simple for them (McNeil *et al.*, 2000). Therefore, studies of both adults and children suggest that the utility of gesture as a scaffold for listeners’ comprehension may depend upon task demands.

This leads to a second consideration, the effect of gesture on learners. High task demands are inherent to learning situations. Not surprisingly, classroom teachers use more gesture when introducing new concepts to their students than when reviewing old concepts (Alibali & Nathan, 2007), and children exposed to new concepts via aligned speech and gesture learn better than those exposed to new concepts via speech alone (Valenzeno, Alibali & Klatzky, 2003). When teachers’ gestures do not align temporally or conceptually with their words, students’ learning suffers (Roth & Bowen, 1999).

Gesture as a contextual support for word learning

In the current study, we were particularly interested in the effect of the speaker’s gestural input on the listener’s performance in an ostensive word-learning context. To ensure that the gestured input provided to the listeners aligned temporally with the word to be learned, we designed a training demonstration in which the examiner gestured while speaking. To provide conceptual alignment between the gesture and the word, we used a symbolic gesture that was iconic, that is, one that readily captured the meaning of the target word.

Two previous investigations suggest the utility of our approach. In one study, Ellis Weismer & Hesketh (1993) taught eight normally developing kindergarteners (and eight children with specific language impairment) three novel words for spatial concepts (e.g. the nonce word/ wug/to mean ‘beside’). These words were modeled repeatedly to describe the spatial location of an alien doll under the pretense that the words were part of the alien’s language. Importantly, for half of the stimulus items, while the examiner modeled the novel words, she also presented an iconic gesture that conveyed the associated spatial meaning. Subsequent comprehension probes (e.g. ‘put Sam wug the box’) revealed a significant advantage for those word–concept pairings that were taught with gestural support.

In a more recent study, Capone & McGregor (2005) employed a similar teaching paradigm. They introduced novel words and their referents within a play setting and, as in Ellis Weismer & Hesketh (1993), the link between the word and referent was emphasized by iconic gesture in some conditions. This study varied from Ellis Weismer & Hesketh (1993) in three ways: the participants were toddlers rather than kindergarteners, the referents were objects rather than spatial locations, and the teaching and testing involved four sessions extended over the course of eleven days rather than a single session. By the fourth session (nine cumulative exposures to the novel words), the children were better able to produce the novel names and were better able to report the function of objects learned in the gesture conditions than in the no-gesture condition.

These two studies were useful first steps towards demonstrating the influence of symbolic gesture on word learning; they also introduced new questions. First, the control condition in both studies was merely an absence of gesture; that is, they involved no symbolic support beyond the words being spoken. Is there something special about gesture as a support for word learning or would any other symbolic support work as well? Second, because both studies included a focus on children’s fast mapping, or their initial appreciation of the word-to-referent link, the introduction of novel words was necessary. This design requirement limited the ecological validity of the studies. Do gestures enrich children’s understanding of

real words, words whose meanings have already begun to emerge from everyday experiences?

In the current study, we pursued these questions by examining toddlers' comprehension of a real target word, the spatial term *under*, before and after training wherein action models and verbal labels were supported by either a gesture, a photograph or by no additional symbol. In the following sections we motivate and expound these methodological decisions.

Is gesture special ?

To build upon the results of Ellis Weismer & Hesketh (1993) and Capone & McGregor (2005), we wished to determine whether gesture as a modality for symbol support is particularly useful or whether any other symbol support would work as well. To do so, we chose to compare the effect of gestured symbols and static symbols on word learning. Given that a spatial term was the target word to be learned, it was essential that both symbols could capture visual–spatial meanings.

Gestures readily express visual–spatial meanings (Alibali *et al.*, 2000; Graham & Argyle, 1975; Majid, Bowerman, Kita, Haun & Levinson, 2004). They can convey location and movement. Both are relevant to the meaning of spatial terms in that such terms serve to relate a landmark object to a trajector object that either is moving or has been moved into a given location (e.g. 'she is hiding the ball under the table' ; 'the ball is under the table'). Moreover, children use spatial terms to express both location and movement (Clark, 2004; Kelly, 2002). For example, two-year-olds will frequently use the word *up* to mean 'pick me up'. Gestured input may serve to reinforce this early appreciation of the semantics of spatial terms.

Gestured symbols are also developmentally appropriate for toddlers. Children exploit the gestural modality for communicative purposes from a very early age. Children who are taught symbolic gestures will use them to communicate before they are able to do so using spoken words (Goodwyn & Acredolo, 1993; McGregor & Capone, 2004). Children who are not directly taught such gestures will create their own and the timing of this attainment roughly parallels the onset of spoken words (Capone & McGregor, 2005). By 1;6, children comprehend requests for objects that are conveyed via gesture (Tomasello, Striano & Rochat, 1999).

Photographs, too, can capture visual–spatial relationships. A photograph of a trajector object together with a landmark object provides a static depiction of their locative relationship. Photographs are also relevant from a developmental standpoint. Even infants aged 0;5 can link a photograph with the object it represents (DeLoache, Strauss & Maynard, 1979) and children aged 2;6 can purposefully use photographs as symbols to guide their search within a three-dimensional context (DeLoache & Burns, 1994). In the context of naming, children aged 1;6 and 2;0 treat pictures (drawings) as symbols for object referents (Preissler & Carey, 2004).

We hypothesized that, because it can convey both the target location and the necessary movement, gesture would be a more useful support for mapping the meaning of *under* than would photographs, symbols that readily convey location but not movement. Given children's early sensitivity to motion (Kellman, 1993), gesture may also serve to elicit more attention to moments of training than still photographs.

Do gestures enrich nascent word knowledge?

Between the ages of 0;9 and 1;2, children begin to demonstrate sensitivity to a range of spatial contrasts (McDonough, Choi & Mandler, 2003). By 1;6, children display heightened sensitivity to spatial categories relevant in their own language community (Choi, McDonough, Bowerman & Mandler, 1999). For children learning English, comprehension of the words *in*, *on* and *under* emerges early: *in* before 2;0, *on* between 2;0 and 2;3, and *under* between 2;0 and 2;9 (Johnston, 1988); however, at such early ages, comprehension remains highly context dependent (Clark, 1973; Wilcox & Palermo, 1974; Rohlfing, 2001; 2006). Specifically, young children are better able to follow instructions that contain spatial terms when those instructions match their understanding of how objects usually function in context (Lloyd, Sinha & Freeman, 1981). A canonical relationship between objects (e.g. ‘put the spoon in the cup’) will elicit better understanding than a non-canonical relationship (e.g. ‘put the spoon on the cup’).

With toddlers aged 1;8 to 2;0 as participants, and *under* as a training target, we were in a position to examine the contribution of gesture to the enrichment of children’s nascent word knowledge. We expected the children to come into the study with some appreciation of spatial categories and, perhaps, some fragile or context-restricted knowledge of *under*. By comparing the children’s ability to follow *under* instructions before and after gesture training, we could examine the extent to which gesture enriched this fragile knowledge. Furthermore, by careful selection of objects, we could examine the extent to which gesture enabled a more context-independent understanding. Specifically, we included object pairs that allowed *under* relationships that were canonical (e.g. ‘girl under umbrella’) and non-canonical (e.g. ‘cup under table’) to determine whether gesture was effective in promoting understanding in both contexts.

We employed a gradient scoring method inspired by Capone & McGregor (2005) to get at degree of knowledge. Specifically, we were not as concerned with whether children knew the meaning of *under* before and after training, but rather whether they needed minimal, moderate or maximal scaffolding to demonstrate that knowledge. Therefore, the specific prediction was that children trained with gesture supports would require less scaffolding after training than before and that their declining need for scaffolding would be more dramatic than that of the children trained with a different symbolic support or than children who received no extra symbolic support.

The current study

To summarize, in the current study, we presented training demonstrations in which objects were placed into *under* relationships and labeled. We compared the utility of gestures and photographs as supports for the mapping of *under* in this training context. To discern the role of symbolic support in general, we also compared the learning of children who received gesture or photograph supports to children who received no extra symbolic support. We selected the term *under* as a target and toddlers between 1;8 and 2;0 as participants because *under* was likely in the realm of the toddlers’ conceptual understanding but was, at the same time, not far along in the learning process. In addition, the choice of *under* extended the focus of previous investigations from novel to real words. We also tested, but did not train, the comprehension of *on* as a means of demonstrating the children’s compliance with the task. We reasoned that, because knowledge of *on* emerges prior to *under*, we could use *on* performance to ensure that each child understood the task and was paying attention to the examiner.

Children's degree of knowledge of *under* was measured in terms of the level of scaffolding they needed to perform *under* instructions before, immediately after and two to three days after training that varied in symbolic support. Given its accessibility to young children and its inherent capacity for capturing movement, we hypothesized that gesture would be superior to photographs and to no symbolic support in promoting enrichment of the spatial term *under*. We also explored the basis for the gesture advantage by determining whether gesture attracted more attention to moments of training than photographs or no symbolic support. Finally, we explored the extent of the gesture advantage by determining whether gesture lessened the children's dependence on physical context (i.e. canonical objects) when processing *under* instructions.

METHOD

Participants

Originally we recruited 49 participants but excluded 9 due to noncompliance ($n = 5$), attrition ($n = 2$), bilingualism ($n = 1$) or a ceiling-level performance on *under* at pretest ($n = 1$). The actual participants were 40 monolingual English-speaking children, 19 boys and 21 girls, between the ages of 1;8 and 2;0. Children were randomly assigned to one of three training groups: model+gesture symbol (+Gesture), model+photograph symbol (+Photo), or model alone (Model Only). The results of three one-way ANOVAs demonstrated that the groups were well matched on chronological age ($F(2, 37) = 0.69, p = 0.51, \eta_p^2 = 0.04$), total number of words ($F(2, 37) = 0.25, p = 0.78, \eta_p^2 = 0.01$), and number of spatial terms ($F(2, 37) = 0.29, p = 0.75, \eta_p^2 = 0.02$), the latter two as reported by parents on the *MacArthur-Bates Communicative Development Words and Sentences Inventory* (Fenson et al., 1993) (see Table 1). Overall children averaged 4.23 ($SD = 4.16$) spatial terms in their expressive vocabularies. According to parent report, one child in the +Gesture group, one child in the +Photo group and two children in the Model Only group had produced the spatial term *under* prior to the study.

Because of random assignment, the training groups were not balanced for numbers of boys and girls. The +Gesture group consisted of 4 boys and 8 girls ; the +Photo group of 10 boys and 5 girls ; and the Model Only group of 6 boys and 7 girls. For this reason, gender was treated as an independent variable in the main analyses.

Stimuli

The stimuli consisted of 10 item pairs (see Table 2 and Appendix A), four for training *under* (dog–roof, girl–umbrella, boat–bridge, book–lamp), four for determining whether trained knowledge of *under* generalized to untrained items (cup–faucet, gift–tree, cup–table, girl–bed), and two for testing compliance with *on* instructions (lettuce–bread, girl–bike). All of the training items were functionally canonical in that the trajector object (e.g. book) would commonly go under the landmark object (e.g. lamp) in everyday environments. The *under* generalization items included two pairs that were canonical (cup–faucet, gift–tree) and two that were non-canonical (cup–table, girl–bed). The latter were included so that we could determine whether training enabled a less context-dependent understanding of *under*. The *on* compliance items were both canonical. Importantly, all object sets, whether canonical or non-canonical, allowed responses other than *under* (or *on*). For example, a child could hold the girl on top of the umbrella, place the cup beside the faucet, or insert the gift in the branches of the tree.

In session one, all children were pretested on one canonical and one non-canonical *under* generalization pair and were post-tested on a different canonical and non-canonical generalization pair with assignment of items to pre- or post-test counterbalanced across

children. In session two, all children were post-tested on all items (four training item pairs, four generalization item pairs, two compliance item pairs).¹

Procedure

Each child was seen in a laboratory setting for two sessions scheduled two or three days apart. The procedures conducted during those sessions were approved by the University of Iowa internal review board for ethical treatment of human subjects. The format of the first session was free play, pretest, training and immediate post-test. The session lasted 20–30 minutes. The format of the second session was free play and delayed post-test. The session lasted 10–15 minutes. Parents were asked not to demonstrate or practice *under* relationships with their children during the days between sessions.

Free play—The purpose of the 5-minute free play was to familiarize the child with the objects and object labels to be used during the experiment proper. The examiner placed all object stimuli on the table and labeled each one once as the child explored and played.

Pretest—The purpose of the pretest was to obtain a baseline measure of *on* and *under* comprehension. *On* was tested purely as a means of determining the child’s compliance with the task. Pretesting of *under* was essential to document each child’s knowledge upon entrance to the study and to ensure that no child who participated already exhibited mastery of *under* (i.e. earned maximum points on both *under* trials as described below).

During pretesting, three item pairs were presented one at a time in random order. Upon each presentation, the examiner (1) gave an instruction that included *on* (in a single instance, e.g. ‘Put the lettuce on the bread’) or *under* (in two instances, e.g. ‘Put the cup under the faucet’). If the child did not comply, the examiner (2) provided a contextual cue that suggested a reason for the trajector to be under (or on) the landmark (e.g. ‘Put the cup under the faucet so we can get a drink’; see Appendix B for a complete list of contextual cues). If the child still did not comply, the examiner (3) demonstrated the expected response and returned the objects to the child saying, ‘Now it’s your turn’, to prompt an imitative response. Scaffolds were always presented in this order and scaffolding of a given item was discontinued as soon as the child answered correctly. Performance was measured on a continuous scale: children were awarded 3 points if they immediately followed the verbal instruction, 2 points if a contextual cue was required, 1 point if an imitative model was required, or 0 points if there was no correct response following these levels of scaffolding. Averaging across items, a mean scaffolded score per child was determined for purposes of data analysis.

Training—For all groups, the examiner presented four item pairs, one at a time. Upon presentation, she labeled each object (‘Here’s the boat and here’s the bridge’), gave the *under* instruction (‘Put the boat under the bridge’) and then immediately completed the targeted relationship while saying, ‘like this’.

For the +Gesture group, the examiner held her right hand over her left hand then moved the right hand under the left as she gave the *under* instruction (see Appendix C).

For the +Photo group, the examiner held up a photograph of the item pair already arranged in the *under* position as she gave the *under* instruction.

¹Note that this use of the stimuli means that the generalization items, though not directly trained, were not wholly unfamiliar. Before testing during session one, each child had seen the generalization items once during free play (see ‘Procedure’). By session two, each child had some experience with the four generalization item pairs in *under* relationships because these had been tested during either the pretest or immediate post-test during session one.

For the Model Only group, the examiner gave the *under* instruction and completed the targeted relationship without any other symbolic support.

For all groups, after the examiner completed the targeted relationship for each item pair, she placed the pair in front of the child and said, ‘Now it’s your turn’ to encourage the child to imitate the placement of the objects in an *under* relationship.

Post-tests—There were two post-tests, one immediately after training and one delayed for two to three days after training. All children participated in both post-tests. The procedure used in the immediate post-test was identical to that used in the pretest except that the stimuli differed. The stimuli were the four item pairs used during training as well as two untrained and untested item pairs for testing *under* generalization and one untrained and untested item pair for testing compliance with *on* instructions. These item pairs and their relevant instructions (*under* or *on*) were presented in random order.

At the delayed post-test, the examiner used identical procedures as the pre- and immediate post-tests to assess compliance with eight *under* instructions (with the four trained item pairs and the four generalization item pairs from session one) and one *on* instruction, presented in random order.

Statistical analysis

With the exceptions noted below, all analyses involved (3) training group×(2) gender×(3) test ANOVAs with repeated measures on the final variable. All significance tests were two-tailed. As an indication of effect size, η_p^2 , or the proportion of the effect+error variance that is attributed to the effect, was computed. Tukey’s HSD for unequal Ns was used for post-hoc testing of between-subject differences and the Bonferroni test was used for post-hoc testing of within-subject differences.

RESULTS

Performance of *on* relationships

Prior to testing the hypothesis of interest, we examined responses to *on* instructions to ensure that the three groups understood and complied with the general protocol. Comparing scaffolded *on* scores via a (3)×(2)×(3) mixed-model ANOVA, we found no main effect of training group ($F(2, 34) = 0.52, p = 0.60, \eta_p^2 = 0.03$), or gender ($F(1, 34) = 2.71, p = 0.11, \eta_p^2 = 0.07$), or training group x gender interaction ($F(2, 34) = 0.15, p = 0.86, \eta_p^2 = 0.008$). There was a main effect of test ($F(2, 68) = 3.95, p = 0.03, \eta_p^2 = 0.10$), such that immediate post-test ($M = 2.25, SE = 0.15$) and delayed post-test ($M = 2.31, SE = 0.16$) performances were significantly better than pretest performance ($M = 1.81, SE = 0.18; p < 0.05$) but not significantly different from each other. These results suggest a ‘warm up’ effect over the course of the first visit. The main effect of test was qualified by a test×group interaction ($F(4, 68) = 3.76, p = 0.008, \eta_p^2 = 0.18$): the warm up was evident for the Model Only and +Photo groups whose average scaffolded scores improved from pre- to post-test on day one by 0.92 and 0.80, respectively, but not for the +Gesture group whose average score was identical at pre- and post-test. There was no test×gender interaction ($F(2, 68) = 0.29, p = 0.75, \eta_p^2 = 0.008$), and no test×training group×gender interaction ($F(4, 68) = 1.36, p = 0.26, \eta_p^2 = 0.07$). The results yield no evidence of drops in compliance over time for any group.

Performance of *under* relationships

Learning of *under* was assessed in two ways. First, we compared the groups’ abilities to perform *under* instructions given the items used by the examiner during the training models.

Second, we compared the groups' abilities to perform *under* instructions with untrained, generalization item pairs.

When mean scaffolded scores for manipulation of trained items were submitted to a (3)×(2)×(3) mixed-model ANOVA, the result was a main effect of training group ($F(2, 34) = 3.89$, $p = 0.03$, $\eta_p^2 = 0.19$), with the Model Only group ($M = 2.07$, $SE = 0.15$) performing significantly better overall than the +Photo group ($M = 1.5$, $SE = 0.15$; $p = 0.03$) but not significantly better than the +Gesture group ($M = 1.6$, $SE = 0.16$). There was no effect of gender ($F(1, 34) = 0.001$, $p = 0.97$, $\eta_p^2 < 0.001$), and no interaction between training group and gender ($F(2, 34) = 0.26$, $p = 0.77$, $\eta_p^2 = 0.02$). There was a main effect of test ($F(2, 68) = 36.23$, $p < 0.001$, $\eta_p^2 = 0.52$), with performance at immediate ($M = 1.93$, $SE = 0.15$) and delayed post-tests ($M = 2.22$, $SE = 0.09$) higher than at pretest ($M = 1.03$, $SE = 0.12$; $p < 0.001$) but not significantly different from each other. There were no interactions between test and group ($F(4, 68) = 1.0$, $p = 0.41$, $\eta_p^2 = 0.06$), test and gender ($F(2, 68) = 1.5$, $p = 0.22$, $\eta_p^2 = 0.04$), or test, group and gender ($F(4, 68) = 0.92$, $p = 0.45$, $\eta_p^2 = 0.05$). The main effect of training group together with a lack of test×group interaction suggests that (1) because of random assignment, the Model Only group came into the study with superior knowledge of *under*² and retained this advantage, and (2) all three groups improved significantly and to a similar degree in their ability to follow *under* instructions given trained item pairs.

A more stringent test of learning is whether the children demonstrated improvements in their comprehension of *under* when given the untrained generalization items (see Table 3). When mean scaffolded scores for generalization items were submitted to a (3)×(2)×(3) mixed-model ANOVA, the result was a main effect of training group ($F(2, 34) = 3.77$, $p = 0.03$, $\eta_p^2 = 0.18$), with the Model Only group ($M = 1.65$, $SE = 0.16$) performing significantly better overall than the +Photo group ($M = 1.09$, $SE = 0.15$; $p = 0.01$) and marginally better than the +Gesture group ($M = 1.21$, $SE = 0.16$; $p = 0.08$). There was no effect of gender ($F(1, 34) = 0.14$, $p = 0.71$, $\eta_p^2 = 0.004$), and no training group×gender interaction ($F(2, 34) = 0.71$, $p = 0.50$, $\eta_p^2 = 0.04$). There was a main effect of test ($F(2, 68) = 13.89$, $p < 0.001$, $\eta_p^2 = 0.29$), with performance at delayed posttest ($M = 1.74$, $SE = 0.10$) higher than at pretest ($M = 1.03$, $SE = 0.12$; $p < 0.001$) and immediate post-test ($M = 1.18$, $SE = 0.14$; $p < 0.001$). Pretest and immediate post-test did not differ significantly. These main effects were qualified by an interaction of test and training group ($F(4, 68) = 2.93$, $p = 0.03$, $\eta_p^2 = 0.15$). This interaction reflected two relevant patterns. First, only the +Photo group demonstrated significant gains from immediate post-test to delayed post-test ($p = 0.004$) but note that this was largely a function of recovering from the dip in performance from pretest to immediate post-test (see Table 2); there was no overall gain for the +Photo group. In fact, the other relevant pattern accounting for the interaction was that, in accord with the primary prediction of the study, only the +Gesture group demonstrated overall gains; their improvement from pretest to delayed post-test was significant ($p = 0.02$; see Table 3). There were no interactions between test and gender ($F(2, 68) = 2.65$, $p = 0.08$, $\eta_p^2 = 0.07$), or test, training group and gender ($F(4, 68) = 0.77$, $p = 0.54$, $\eta_p^2 = 0.05$).

Exploration of the gesture effect—To further explore the gesture advantage, we conducted three analyses. Specifically, we sought to better understand the basis, extent and time course of the gesture effect.

²Recall from Table 1 that the Model Only group did not have larger spatial term vocabularies than the other groups upon entrance to the study; however, this does not negate the possibility that they had a stronger understanding of the meaning of *under*. Spatial term vocabulary was estimated via parents' reports of expressive knowledge whereas the probes used in all test phases of the study itself involved observation of receptive knowledge.

To explore the basis of the gesture effect, we compared the attentiveness of the groups. We reviewed videotapes to tally the number of times that each child imitated the examiner's four training models to test the hypothesis that the examiner's gesture was effective because it attracted attention to, and participation in, moments of training. Useable videos were available for 10 children in the Model Only group, 9 children in the +Gesture group, but because of a damaged digital tape, only 5 children in the +Photo group. Based on this partial dataset, all children looked at least briefly at each of the examiner's models and all groups imitated, on average, more than half the time (+Gesture: $M = 2.33$, $SD = 1.22$; +Photo: $M = 2.60$, $SD = 1.14$; Model Only: $M = 3.30$, $SD = 0.67$). A t -test comparing the rate of imitations in the +Gesture group to that of the +Photo and Model Only groups combined revealed no significant attentional advantages for the +Gesture group ($t = 1.7$, $df = 22$, $p = 0.10$, Cohen's $d = -0.72$). In fact, the non-significant trend was such that the children in the +Gesture group were slightly less engaged by the training.

We also explored the extent to which the learning of the +Gesture group varied with object context. To do so, we categorized the scores of the +Gesture group in response to the generalization items at each test (Figure 1). In this descriptive analysis, a canonicity effect was apparent in that the canonical item pair *cup-faucet* consistently elicited the highest level of performance whereas the non-canonical item pair *cup-table* elicited the poorest performance. Improvement from pretest to delayed post-test was apparent for all items except *cup-table*. Degree of improvement on non-canonical *girl-bed* (and on canonical *gift-tree*) was as large as or larger than the degree of improvement for *cup-faucet*. Therefore, gesture supports did improve comprehension in one non-canonical context (*girl-bed*) but not in another (*cup-table*).

Finally we explored the delay in emergence of the gesture advantage. As a group, there was no significant improvement from pretest to immediate post-test; the advantage for the +Gesture group was revealed in the overall gain from pretest to delayed post-test. Did these mean scores mask individual differences such that some children demonstrated immediate gains in response to gesture-enhanced training but others (the majority) demonstrated gains only at delayed post-test? Alternatively, did children who demonstrate modest gains on the immediate post-test build on those gains for a more impressive performance at delayed post-test? A strong correlation obtained between short-term and overall gain scores for the +Gesture group ($r(12) = 0.87$, $p < 0.001$) supported the latter explanation. The extent to which a child demonstrated short-term benefits from the training demonstrations and the extent to which that child improved overall were related.

DISCUSSION

Gestured input enhances communication on the part of preverbal children (Acredolo & Goodwyn, 1988); it strengthens early language development of verbal children (Goodwyn, Acredolo & Brown, 2000), and, more specifically, it facilitates their initial mapping (Capone & McGregor, 2005; Ellis Weismer & Hesketh, 1993) and subsequent enrichment of novel words (Capone & McGregor, 2005). This study reveals an additional benefit of gestured input for verbal children: Gesture served to enrich their knowledge of a real word. Children who received training supplemented with symbolic gesture demonstrated more robust long-term gains in word knowledge, not only as compared to children who received the same training without symbolic support, but also as compared to children whose training was supplemented by symbolic photographs. In the following sections, we interpret the nature of the gesture effect and present some hypotheses about the basis of the gesture advantage.

The nature of the gesture effect

The children in this study had almost certainly been exposed to *under* prior to the study. On average, they came in with some receptive knowledge of the word; by chance, the +Model Group demonstrated more knowledge than the other two groups upon recruitment. They retained this advantage but they did not learn to a greater degree than children in the other groups during the course of the study. On average, the children in all three groups were good word learners. When given the items that the examiner had used during training, all of which allowed for canonical *under* relationships, all three groups of children improved from pre- to immediate post-test and maintained that improvement at delayed post-test. This improvement may have reflected, in part, the warm-up characteristic of performance with untrained *on*; however, this is not likely the complete explanation as children in the +Gesture group did not demonstrate warm-up with *on* but did, like their peers in the other two groups, demonstrate improvement with *under*. It seems that the children gained some additional knowledge of *under* and they did so whether or not they received symbolic support in the form of gesture or photographs.

The gesture advantage was revealed by the children's ability to follow *under* instructions given the untrained generalization items. With these items, only the children in the +Gesture group, those whose training demonstrations had been supplemented by symbolic gesture, demonstrated long-term gains. The results are reminiscent of Rohlfing (2006) who compared children learning *under* with or without the benefit of a verbal contrast (e.g. 'Look the dog is under the table, not on!'). She found no effect of the contrast when the children were tested with the training materials but a contrast advantage emerged when they were tested with untrained materials. We maintain that the differential sensitivity of tests employing trained and untrained items reflects the relationship between the need for scaffolding and the robustness of knowledge. Compared to the untrained generalization items, trained items were better scaffolds because the children had more experience with them and, in particular, more exposure to them in *under* relationships.³ Even modest gains in semantic knowledge could be elicited when scaffolded by trained items; a more robust knowledge was required to perform with the reduced scaffolding provided by untrained items. Only with the semantic enrichment of a contrasting term, in Rohlfing (2006), or a symbolic gesture, in the current study, did more robust learning take place.

Part of this robustness was reflected in the fact that, after training, the +Gesture group became somewhat less dependent on the context provided by canonical objects. Two non-canonical item pairs were tested and gains on one, *girl-bed*, were impressive. However, there was no gain on the non-canonical item pair *cup-table*. This difference may reflect the extent to which these item pairs were truly non-canonical. For children in particular, hiding under the bed may be familiar and therefore somewhat canonical, despite being less common than lying on the bed. Moreover, cups normally go on tables and flat surfaces such as tables invite *on* substitutions (Clark, 1973), a common error among the children in the current study. Gesture did not help the children to overcome these particularly powerful affordances for the spatial relationship *on*.

³Recall that, by session two, the children did have some chances to experience the untrained generalization items in *under* relationships. Each item pair had been tested during session one, which means that the child had either followed an instruction to place the items, one under the other, or, if unable to do so, had observed the examiner do so in an effort to scaffold the child's response. The child may well have learned something about the concept of *under* during the test. The important points here are that : (a) exposure to the trained items was greater than exposure to the generalization items, involving not only free play and testing in session one but also the training exposure itself ; and (b) the test exposures were constant for the three training groups yet the learning outcomes differed.

The time course of the gesture effect

The gesture advantage was revealed two to three days after training. Short-term gains were more subtle and did not reach statistical significance. One mundane explanation for this delay is that, by the end of the first visit, the children were fatigued or bored and, as a result, they did not perform at their best on the immediate post-test. This argument suffers in light of the *on* and trained *under* data. If the children had been fatigued or bored with the task in general, their performance with *on* instructions during the immediate post-test should have suffered, but this was not the case. All three groups followed *on* instructions as well or better at post-tests than at pretest. Also, improvement with *under* instructions from pre- to immediate post-test given trained items suggests the children were not bored by *under* instructions in particular.

Another possibility is that the children in the +Gesture group learned the *under* instruction outside of the laboratory during the interval between immediate and delayed post-tests. Parents were explicitly instructed not to practice *under* instructions with their children. It may be that some parents did not follow this instruction but there is no reason to suspect that parents of children in the +Gesture group would be less compliant than parents of children in the other groups. Of course *under* is a real word and there was no way to control the children's exposure in naturalistic contexts to the word during the interval between tests but, again, the possibility of inadvertent exposure was equal across groups. The lack of significant improvement in the comparison groups weakens the possibility that practice or exposure outside of the lab accounts for the results.

A more interesting possibility is that, over the interval between immediate and delayed post-tests, the information gleaned from the gesture-enhanced training consolidated. Consolidation involves stabilization and enhancement of information in memory over time in the absence of additional exposure to that information (Walker, 2005). The positive correlation between gains at immediate post-test and gains at delayed post-test is consistent with the hypothesis that consolidation was at play. Consolidation of newly learned words has been documented in adult learners (Dumay & Gaskell, 2007) and, though not examined directly, patterns that hint at consolidation effects in children's word learning do exist (Rice, Oetting, Marquis, Bode & Pae, 1994; Storkel, 2001). Future studies aimed at testing the consolidation hypothesis should ensure that there is no exposure to the newly learned word between gesture-enhanced teaching and delayed test.

The utility of gesture to the mapping process

There could be several reasons for the success of the gesture-supported teaching. One possibility is that gestures are interesting, and thus draw more attention to moments of training. However, we can rule out this possibility because all groups of children paid close attention to the examiner during the training demonstrations; the +Gesture group was not at all superior in this regard. This does not mean that gesture never serves to enhance attention to ostensive naming episodes nor that such enhancement never serves to facilitate learning. For example, in a study of the effect of deictic gestures on the fast mapping behaviors of children aged 2;4 to 2;7, there was a positive relationship between attention and learning such that the more a particular type of gesture drew the child's attention to the target at the time of labeling, the better the child's performance at posttest (Booth, McGregor & Rohlfing, 2008). Perhaps gesture as an inducement to attend may be particularly useful for promoting fast mapping but less useful in promoting subsequent learning wherein the goal is not to infer a word to referent link but, rather, to enhance knowledge of the word meaning.

A more likely explanation for the gesture effect in the current study is one suggested by Goldin-Meadow (2000). She posits that gesture serves to minimize cognitive load during

moments of language processing. Because the gesture used here was iconic and because it was presented in temporal contiguity with the word, it externalized a meaningful aspect of the referent in the visual world. By making that meaning more obvious, gesture may free cognitive–linguistic resources for processing the word itself and, perhaps, the other lexical and syntactic elements involved. This should be especially true in the current study. Because we recruited toddlers who were very much in the throes of acquiring spatial meanings, task demands associated with *under* instructions were high, and therefore the scaffolding provided by the gesture should have been particularly useful.

Although the comparison between gesture and photographs as a support for the enrichment of *under* favored gesture, there remain some gaps in our understanding of the utility of gesture to word learning. If one of the benefits of gesture was, as we have argued, the expression of movement, then other expressions of movement might be equally helpful. Video animations involving movement of inanimate objects into relevant spatial relationships would provide a suitable test.⁴ Also, certain word meanings might be better conveyed via photographs than gestures. Experiments designed to teach words whose meanings are less tied to movement would provide a suitable test. Finally, there might be certain learners whose learning styles are not highly amenable to gestured input. The frequent use of photographs or drawings in clinical language intervention settings is suggestive. Quill (1997) argues that pictures may be especially effective for learners, such as many with autism spectrum disorders, who have stronger visual than auditory perception and who are better able to attend to stable than fleeting referents. The efficacy of gesture and still images as learning aides for these populations should be tested.

SUMMARY AND CONCLUSIONS

Input that included symbolic gesture enriched children’s understanding of the spatial term *under* and did so to a greater extent than input that included symbolic photographs (or no supplemental symbols). Only measures employing untrained materials and delayed testing were sensitive to the gesture advantage.

We conclude that gestured input promoted more robust knowledge of the meaning of *under*, knowledge that was less tied to contextual familiarity and more prone to consolidation. Gesture may have been particularly effective because it reduced cognitive load during training while emphasizing the target location as well as the movement inherent in the meaning of *under*. Word learning is a gradual process, one that progresses from emergence to mastery as the learner repeatedly attaches meaning to the word in multiple contexts (Bloom, 2000). This study demonstrates that symbolic gestures, as useful contextual supports for children’s comprehension of words in the moment, are ultimately useful supports for word learning.

References

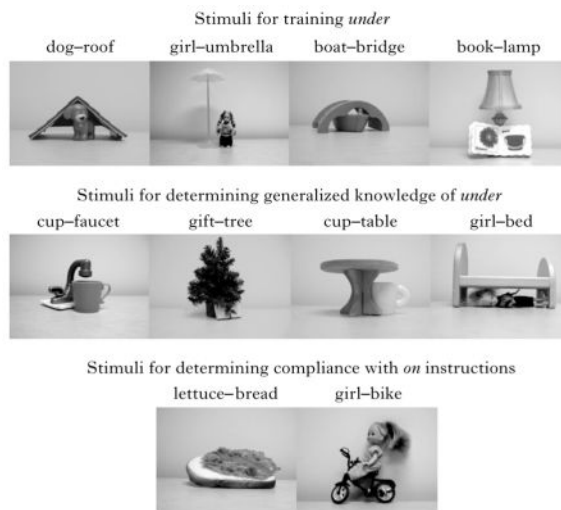
- Acredolo LP, Goodwyn SW. Symbolic gesturing in normal infants. *Child Development*. 1988; 59:450–66. [PubMed: 2452052]
- Alibali MW, Kita S, Young AJ. Gesture and the process of speech production : We think, therefore we gesture. *Language and Cognitive Processes*. 2000; 15:593–613.
- Alibali, MW.; Nathan, MJ. Teachers’ gestures as a means of scaffolding students’ understanding: Evidence from an early algebra lesson. In: Goldman, R.; Pea, R.; Barron, BJ.; Derry, S., editors. *Video research in the learning science*. Mahwah, NJ: Erlbaum; 2007. p. 349-65.
- Bloom, P. *How children learn the meanings of words*. Cambridge, MA: MIT Press; 2000.

⁴We thank an anonymous reviewer for this suggestion.

- Booth AE, McGregor KK, Rohlfing K. Socio-pragmatics and attention : Contributions to gesturally guided word learning in toddlers. *Language Learning and Development*. 2008; 4:179–202.
- Capone NC, McGregor KK. The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, & Hearing Research*. 2005; 48:1468–80.
- Choi S, McDonough L, Bowerman M, Mandler JM. Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development*. 1999; 14:242–68.
- Clark EV. Non-linguistic strategies and the acquisition of word meanings. *Cognition*. 1973; 3:161–82.
- Clark EV. How language acquisition builds on cognitive development. *TRENDS in Cognitive Sciences*. 2004; 8:472–78. [PubMed: 15450512]
- DeLoache JS, Burns NM. Early understanding of the representational function of pictures. *Cognition*. 1994; 52:83–110. [PubMed: 7924202]
- DeLoache JS, Strauss MS, Maynard J. Picture perception in infancy. *Infant Behavior and Development*. 1979; 2:77–89.
- Dumay N, Gaskell MG. Sleep-associated changes in the mental representation of spoken words. *Psychological Science*. 2007; 18:35–39. [PubMed: 17362375]
- Ellis Weismer S, Hesketh L. The influence of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech and Hearing Research*. 1993; 36:1013–26. [PubMed: 8246467]
- Fenson, L.; Dale, PS.; Reznick, JS.; Thal, D.; Bates, E.; Hartung, JP.; Pethick, ST.; Reilly, JS. *The MacArthur Communicative Development Inventories (CDI)*. San Diego, CA: Singular Publishing Group, Inc; 1993.
- Goldin-Meadow S. Beyond words: The importance of gesture to researchers and learners. *Child Development*. 2000; 71:231–39. [PubMed: 10836578]
- Goodwyn SW, Acredolo LP. Symbolic gesture versus word: Is there a modality advantage for onset of symbol use? *Child Development*. 1993; 64:688–701. [PubMed: 8339689]
- Goodwyn SW, Acredolo LP, Brown CA. Impact of symbolic gesturing on early language development. *Journal of Nonverbal behavior*. 2000; 24:81–103.
- Graham JA, Argyle M. A cross-cultural study of the communication of extraverbal meaning by gestures. *International Journal of Psychology*. 1975; 10:57–67.
- Johnston, JR. Children's verbal representations of spatial location. In: Stiles-, Davies J.; Kritchevsky, M.; Bellugi, U., editors. *Spatial cognition*. Hillsdale, NJ: Lawrence Erlbaum; 1988. p. 195-205.
- Kellman, PJ. Kinematic foundations of infant visual perception. In: Granrud, C., editor. *Visual perception and cognition in infancy*. Carnegie Mellon Symposia on Cognition. Hillsdale, NJ: Erlbaum; 1993. p. 121-73.
- Kelly, BF. 'Well you can't put your swimsuit on top of your pants!': Child-mother uses of *in* and *on* in spontaneous conversation. 2002. Retrieved 10 September 2008, from http://eslipublications.stanford.edu/CLRF/2002/Pp_69-78_Kelly.pdf
- Kendon A. Do gestures communicate? A review. *Research on Language and Social Interaction*. 1994; 27:175–200.
- Lloyd SE, Sinha CG, Freeman NH. Spatial reference systems and the canonicity effect in infant search. *Journal of Experimental Child Psychology*. 1981; 32:1–10.
- Majid A, Bowerman M, Kita S, Haun DBM, Levinson SC. Can language restructure cognition? The case for space. *TRENDS in Cognitive Sciences*. 2004; 8:108–114. [PubMed: 15301750]
- McDonough L, Choi S, Mandler JM. Understanding spatial relations : Flexible infants, lexical adults. *Cognitive Psychology*. 2003; 46:229–59. [PubMed: 12694694]
- McGregor KK, Capone NC. Genetic and environment interactions in determining the early lexicon : Evidence from a set of tri-zygotic quadruplets. *Journal of Child Language*. 2004; 31:311–37. [PubMed: 15495843]
- McNeil N, Alibali M, Evans J. The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior*. 2000; 24:131–50.
- Preissler MA, Carey S. Do both pictures and words function as symbols for 18- and 24-month-old children? *Journal of Cognition & Development*. 2004; 5:185–212.

- Quill KA. Instructional considerations for young children with autism: The rationale for visually cued instruction. *Journal of Autism and Developmental Disorders*. 1997; 27:697–714. [PubMed: 9455729]
- Rice M, Oetting J, Marquis J, Bode J, Pae S. Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 1994; 37:106–122.
- Rohlfing, KJ. No preposition required. The role of prepositions for the understanding of spatial relations in language acquisition. In: Pütz, M.; Niemeier, S.; Dirven, R., editors. *Applied cognitive linguistics I : Theory and language acquisition*. Berlin: Mouton de Gruyter; 2001. p. 230-47.
- Rohlfing KJ. Facilitating the acquisition of UNDER by means of IN and ON – a training study in Polish. *Journal of Child Language*. 2006; 33:51–69. [PubMed: 16566320]
- Roth WM, Bowen GM. Decalages in talk and gesture : Visual and verbal semiotics of ecology lectures. *Linguistics & Education*. 1999; 10:335–58.
- Storkel HL. Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*. 2001; 44:1321–37.
- Tomasello M, Striano T, Rochat P. Do young children use objects as symbols? *British Journal of Developmental Psychology*. 1999; 17:563–84.
- Valenzeno L, Alibali MW, Klatzky R. Teachers’ gestures facilitate students’ learning : A lesson in symmetry. *Contemporary Educational Psychology*. 2003; 28:187–204.
- Walker MP. A refined model of sleep and the time course of memory formation. *Behavioral and Brain Sciences*. 2005; 28:51–64. discussion 64–104. [PubMed: 16047457]
- Wilcox S, Palermo DS. ‘In’, ‘on’, and ‘under’ revisited. *Cognition*. 1974; 3:245–54.

APPENDIX A



APPENDIX B

Contextual cues used to scaffold children’s understanding of *under* instructions at test

dog–roof	Put the doggie under the roof because he wants to hide.
girl–umbrella	Put the girl under the umbrella because it’s raining.
boat–bridge	Put the boat under the bridge because it’s stormy.
book–lamp	Put the book under the lamp so we can read it.
cup–faucet	Put the cup under the faucet so we can get a drink.
gift–tree	Put the gift under the tree because we’re not ready to open it.

- cup–table Put the cup under the table so we can feed the dog.
- girl–bed Put the girl under the bed because she wants to hide.
- lettuce–bread Put the lettuce on the bread so we can make a sandwich.
- girl–bike Put the girl on the bike because she wants to ride.

APPENDIX C

An example of training for the +Gesture group

An example of training for the +Gesture group



'Here's the boat and here's the bridge.'



'Put the boat under the bridge'



'Like this.'

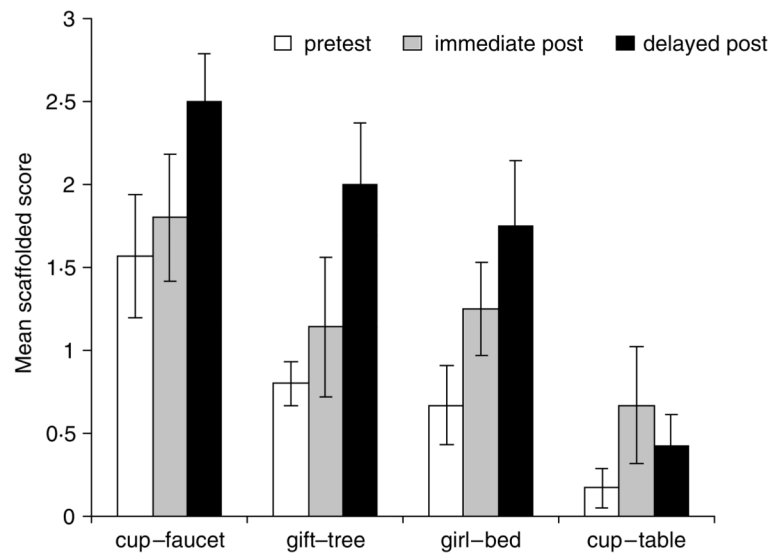


Fig. 1. Mean scaffolded scores of children from the +Gesture group by generalization item pairs and test.
NOTE: Each child responded to two generalization item pairs at pretest, two others at immediate post-test, and to all four at delayed post-test.

TABLE 1

Characteristics of participants per group expressed as means (and standard deviations)

Group	Age	Total words	Total spatial terms
+Gesture	21.00 (1.54)	145.75 (98.86)	4.33 (4.90)
+Photo	21.26 (1.38)	175.61 (84.82)	3.34 (2.46)
Model Only	20.68 (0.95)	165.84 (142.45)	4.24 (4.01)

TABLE 2

Item pairs, their purpose in this study, and whether they hold canonical or non-canonical spatial relationships

<i>Item</i>	<i>Purpose</i>			<i>Type of relationship</i>	
	Training	Generalization	Compliance	Canonical	Non-canonical
dog-roof	X			X	
girl-umbrella	X			X	
boat-bridge	X			X	
book-lamp	X			X	
cup-faucet		X		X	
gift-tree		X		X	
cup-table		X			X
girl-bed		X			X
lettuce-bread			X	X	
girl-bike			X	X	

TABLE 3

Patterns of learning presented as mean scaffolded scores and mean gains (with standard errors) for generalization items

Group	Pretest	Immediate Post-test	Delayed Post-test	Short-term gain (immediate post-pretest)	Intermediate gain (delayed post-intermediate post)	Overall gain (delayed post-pretest)
+Gesture	0.79 (0.23)	1.08 (0.25)	1.75 (0.18)	0.29 (0.33)	0.67 (0.24)	0.88* (0.20)
+Photo	1.04 (0.21)	0.64 (0.23)	1.59 (0.16)	-0.40 (0.30)	0.95** (0.21)	0.52 (0.18)
Model Only	1.24 (0.21)	1.83 (0.25)	1.90 (0.18)	0.58 (0.32)	0.05 (0.23)	0.56 (0.19)

** $p = 0.004$;

* $p = 0.02$.