Developmental Science

Developmental Science 12:1 (2009), pp 81–87 DOI: 10.1111/j.1467-7687.2008.00735.x

PAPER

Co-speech gesture as input in verb learning

Whitney Goodrich and Carla L. Hudson Kam

Department of Psychology, University of California, Berkeley, USA

Abstract

People gesture a great deal when speaking, and research has shown that listeners can interpret the information contained in gesture. The current research examines whether learners can also use co-speech gesture to inform language learning. Specifically, we examine whether listeners can use information contained in an iconic gesture to assign meaning to a novel verb form. Two experiments demonstrate that adults and 2-, 3-, and 4-year-old children can infer the meaning of novel intransitive verbs from gestures when no other source of information is present. The findings support the idea that gesture might be a source of input available to language learners.

Introduction

People have traditionally thought of the input for language acquisition in terms of the speech stream itself. As a result, the vast majority of language development studies have relied on written transcripts of the speech young children hear around them. Recently, however, researchers have begun to focus on additional cues in the environment that, while available to children, are not apparent in the transcripts. For example, eye gaze can be used as a cue to the speaker's communicative intention, and help a child infer the meaning of a word (Baldwin, 2000). This study explores an additional potential source of input – gestures that co-occur with speech – and asks whether learners can use the gestures other people produce to learn the meanings of novel verbs.

Gestures present an intriguing possible source of information, as they are highly communicative, like speech, but in many ways are part of the communicative context, like eye gaze. They are also abundant in natural speech. Around the world, wherever you see people talking you will see people moving their hands (Goldin-Meadow, 2003; Kendon, 2004). These hand movements are far from random: gestures contain information related to the speech they occur with, sometimes matching the content of speech, other times providing additional information (McNeill, 1992). Importantly, there is evidence that mothers gesture in similar ways when speaking to their children. In a study examining Italian mother–child dyads, Iverson, Capirci, Longobardi and Caselli (1999) found that mothers' gestures tended to co-occur with speech. Moreover, the gestures tended to relate to the immediate linguistic context, often highlighting or attracting attention to particular objects that were currently the focus of conversation (see also Zukow-Goldring & Ferko, 1994).

Various lines of evidence demonstrate that listeners can take advantage of the information present in gesture. If a speaker says 'He went down the street' while producing a rolling gesture with their index finger, for example, an adult listener might interpret this to mean 'He went *rolling* down the street' (Bavelas, 1994). Moreover, adults are often not consciously aware of what modality an idea was communicated in, as they will integrate a speakers' speech and gesture into a single unified interpretation (Cassell, McNeill & McCullough, 1998). Young children can also do this. For example, if a speaker says 'push' while pointing at a ball, children will push the ball (Morford & Goldin-Meadow, 1992).

In general, studies exploring learning in children have focused on gestures with noun-like referential content. Namy and Waxman (1998) investigated whether young children would accept a hand sign without accompanying speech as a category name. They found that 18-month-old infants were willing to do so, while 26-month-old infants did not, unless given explicit training. These findings suggest that young children will initially accept either gesture or speech as a symbolic label, while older children have learned the social convention of relying on spoken words as labels (Namy & Waxman, 1998). More recent studies have demonstrated a U-shaped trajectory, showing that 18-month-old infants and 4-year-old children were willing to map both arbitrary and iconic gestures onto a referent, while 26-month-old infants succeeded only at an iconic gesture mapping (Namy, Campbell & Tomasello, 2004). While these studies have exciting implications for symbolic development in young children, they do not necessarily inform us about how children might use gestures that co-occur with (as opposed to replace) speech. They do, however, demonstrate one of the necessary preconditions, that young children are

Address for correspondence: Whitney Goodrich, University of California, Berkeley, Department of Psychology, Room 3210, Tolman Hall #1650, Berkeley, CA 94709, USA; e-mail: wgoodrich@berkeley.edu

able to link abstract gestural representations (at least of an object) with the real-world features represented in the gesture.

Clearly, then, listeners can process the meaning present in gestures, but can they use it to learn something about the speech itself? In this study we focused on the information available in iconic gestures, and asked whether it can help learners acquire the meanings of verbs. Verb learning is generally considered to be a difficult task (Gentner, 1981; Gleitman & Gleitman, 1992). Children rarely hear speech that directly corresponds to actions occurring in the environment. For example, when a mother opens a door, she seldom says 'I am opening the door!' but rather, 'Hello!' (Gleitman & Gleitman, 1992). However, motion verbs are often accompanied by iconic gestures (Hadar & Krauss, 1999). Iconic gestures are concrete, relatively transparent representations, such as tracing a steep curvy road while talking about walking down the famous Lombard Street in San Francisco, or holding your hands wide apart while describing the huge fish you just caught. Despite the iconic nature of such gestures, they are not simply pantomime. They require some inference on the part of the listener to be interpreted, and their meaning is to some degree dependent upon the context in which they occur, especially the accompanying speech (McNeill, 1992).

This study examines whether iconic gestures can help a learner disambiguate the meaning of a new verb when no other cues are available. We presented listeners with a novel verb form, and assessed their interpretation of the novel word using a forced choice task. Sometimes the speech in which the novel form was presented was accompanied by an iconic gesture related to possible verb meanings and sometimes it was not. Crucially, we did not provide any information about the meaning of the verb except for the information contained in the co-speech gesture. The question was whether listeners would use the motion information in the gesture to interpret the novel verb form. Note that we were not testing whether children can learn to use a gesture as a symbolic label for a novel action. Clearly, children learning signed languages can do this, and previous research by Namy and colleagues demonstrates that children learning a spoken language are willing to map arbitrary and iconic symbols to nominal referents, regardless of modality (Namy & Waxman, 1998; Namy *et al.*, 2004). We are quite specifically asking whether the information in gestures can help learners acquire knowledge about the speech it accompanies.

We examined this in children of various ages, as well as adults. The youngest children tested were 2 years of age. Two-year-olds were selected because they are old enough to understand the task, yet still at an age when they are rapidly learning new verbs (Tomasello, 1992). Moreover, there is evidence that 2-year-olds can integrate information from speech and gesture (Morford & Goldin-Meadow, 1992; Namy & Waxman, 1998; Namy *et al.*, 2004). However, from an information processing perspective our task is more difficult than Namy's, as the child has to integrate multiple *simultaneous* sources of information, so we also looked at 3- and 4-year-old children, based on literature suggesting that integrating information from speech and gesture can sometimes be challenging for young children (Kelly & Church, 1988; McNeil, Alibali & Evans, 2000).[1]

Experiment 1

Method

Participants

Ten participants per age per condition were tested, for a total of 30 in each age group. Participants included 2- (mean = 2;7, range: 2;3–2;9), 3- (mean = 3;5, range: 3;1–3;8), and 4-year-old children (mean = 4;5, range: 4;0–4;11), and adults. Adult participants were undergraduates at the University of California, Berkeley, who received course credit for participation. The children were recruited from child-care centers on campus. All participants were native English speakers with no known hearing problems. There were approximately equal numbers of male and female participants.

Design and procedure

The participants' task was to learn four novel intransitive verbs that referred to novel actions performed by unfamiliar toys. Participants heard the following instructions: 'This is my friend Sam. Sam has some funny toys that do funny things. But he likes some of his toys better than others. I'm going to show you Sam's toys and you get to pick his favorite, okay?' On each trial two toys and the actions they performed were demonstrated one at a time, and then the experimenter said, 'Sam really likes to (*novel verb*). Which toy lets Sam go (*novel verb*)-ing?' Participants were required to respond by pointing to, touching, or otherwise indicating their selection.

There were four trials, each with different actions and verbs. The four verbs were *sib*, *blip*, *gern*, and *flim*. The four toy pairs (and their actions) were: (1) a ramp that the puppet rolled down in a tube and a spring that the puppet sat on and bounced back and forth, (2) a stick that bounced the puppet up and down and a catapult that flung the puppet into the air, (3) a stick that arched the puppet up and down in a fan-like motion, and a

[1] There is some previous work showing that (old) 2-year-olds can integrate information from speech and descriptive gestures to learn novel adjectives (although not at ceiling levels of performance; O'Neill, Topolovec & Stern-Cavalcante, 2002). However, descriptive gestures are actual interactions with objects. The co-speech gestures we are exploring are quite different; they are abstract representations that require some inference on the part of the listener if they are to be interpreted. This makes them cognitively more complex, and thus, we might expect poor performance by the younger children.

stick that pulled the puppet along a wavy path, and (4) a round turntable that spun the puppet around and a ramp the puppet was pulled up. The pairing of toy/action and verb was counterbalanced across participants, such that the correct response for one half of the participants was the incorrect response for the other half. We also counterbalanced the order of presentation of the sets of toys and avoided having the same side always being associated with the correct answer.

Participants were randomly assigned to one of three conditions. In the iconic-gesture condition (IG), the statement was accompanied by a gesture that could be used to infer the meaning of the verb. The gesture was produced in synchrony with the novel verb in speech, and occurred both times the experimenter stated the verb. Importantly, this gesture was the only source of information available about the verb's meaning. For example, for the set including the catapult that flung the puppet into the air and the stick that bounced the puppet up and down, the gesture was either an index finger tracing the upward trajectory of the catapult, or an index finger bouncing up and down. Pictures of each toy and its accompanying gesture in the IG condition are shown in Figure 1.

We did not intentionally control for the nature of the gesture (for example whether the gesture encoded the manner or path of the action) as they were intended to be as naturalistic as possible – that is, what a speaker might naturally produce when talking about that action. However, the gesture used with each individual toy was always the same. The toys were matched for 'interestingness' to avoid children having a preference for one over the other. One side effect of this was that the gestures in each pair were rather naturally matched for dynamism. And a post-hoc analysis of the gestures reveals that they were also fairly well matched in terms of the information they encoded: both gestures encoded path only for three of the pairs, while in the fourth both gestures encoded manner and path (see Figure 1).[2] The gestures were somewhat more exaggerated than typical gestures, but we have evidence from an ongoing study in our lab that naïve adult participants use very similar gestures when describing the actions performed by the toys.

There were two control conditions. In the other-gesture condition (OG), the experimenter produced an interactive gesture (see e.g. Bavelas, 1994) that could not be used to disambiguate the meaning of the verb. The four gestures were as follows: (1) the experimenter pointed at the child while saying, 'can you show me which toy?', (2) the experimenter flipped up her hands while saying 'which toy', (3) the experimenter shrugged her shoulders while saying 'which toy', and (4) the experimenter shrugged her shoulders and flipped up her hands at the same time while saying, 'which toy'. The order of the gestures was the same for all participants in this condition. This condition was included to control for the fact that no gesture at all might seem odd for participants. The second control condition was the no-gesture condition (ØG). In this condition, the experimenter kept her hands in her lap while speaking.

Results and discussion

In the IG condition, participants' responses were scored as correct if they selected the toy that matched the iconic gesture. This was not possible for participants in the control conditions, as they did not see any iconic gestures, and thus there was no inherently correct answer. To get around this, each participant in the ØG and OG conditions was randomly assigned to one of the two IG condition 'answer keys' and their response was considered correct if it would have been correct for the IG participant. Thus, each toy in a set was the correct response for half of the control participants, as with the IG participants.

Participants' performance is shown in Figure 2. From the figure it appears that participants in the IG condition performed differently than the other conditions, that is, they seem to have used the gesture to interpret the verb. This impression is borne out by the statistical analysis: an ANOVA revealed a main effect of Condition ($F(2, 108) = 36.413$, $p < .001$). The main effect of Age approached significance ($F(3, 108) = 2.233$, $p = .08$); however, there was no Age by Condition interaction ($F(6, 108) = 1.619$, $p = .149$). The Age effect appears to be driven primarily by the adults' perfect performance in the IG condition. When the adults are excluded from the analysis, Age is not significant ($F(2, 81) = 1.368$, $p = .261$), while the main effect of Condition remains ($F(2, 81) = 13.728$, $p < .001$).

We then went on to ask if participants selected the correct toy (i.e. the toy that performed the action matching the experimenter's gesture) at a level significantly above chance (= 50%). Given that Age approached significance, we performed separate t-tests for each condition within each age group. Performance was significantly above chance for the participants in the IG condition for all age groups (2s: $t(9) = 3.25$, $p = .004$; 3s: $t(9) = 6.0$, $p = .0001$; 4s: $t(9) = 4.99$, $p = .0003$; adults' performance was at ceiling with no variation). This was not the case for participants in the two control conditions; they were all at chance. This is not surprising as in the control conditions there was no information available in the speech, gesture, or social context about the verbs' meanings.

The previous analyses were done with group means which could possibly hide some age-related variability. In particular, it could be the case that in some age groups all participants were equally sensitive to the iconic gestures whereas in others some participants were and some weren't. To investigate this, we examined the number of participants within the IG condition that got 0–2 correct

[2] We coded the gesture that accompanied the Spring toy as containing manner as well as path because of the upright orientation of the arm (and hand). The two components can logically be separated; if the arm were at a different angle, parallel with the floor for instance, but the same motion was produced, it would be manner alone.

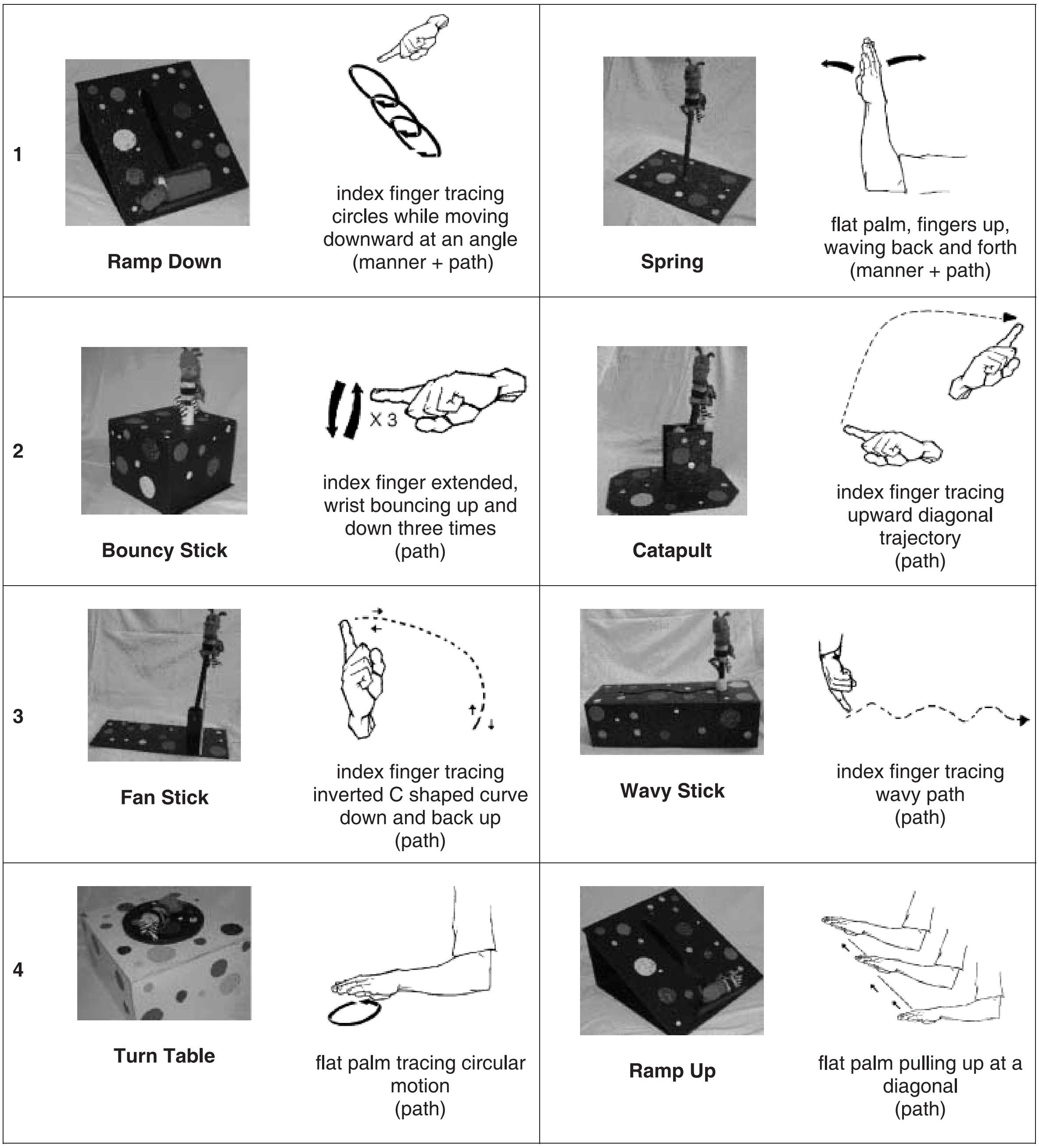

Figure 1 *Picture of toys and accompanying iconic gestures.*

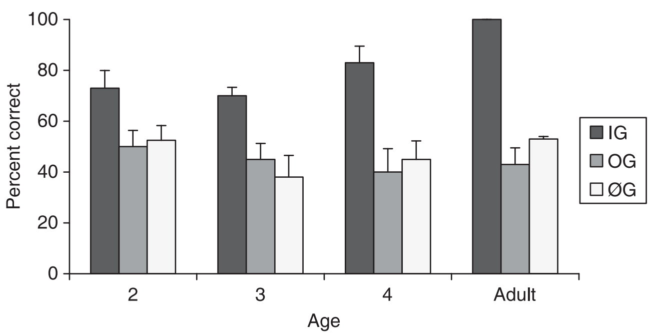

Figure 2 *Percent correct toy choice by age and input type.*

(at or below chance) or 3–4 correct (above chance) by age group. Although there was some variation by age group (six 2-year-olds, eight 3-year-olds, eight 4-year-olds, and 10 adults selected the correct toy 3–4 times out of 4), the differences were not significant, $\chi^2(1, N = 40) = 1.875$, $p = .17$. By contrast, the number of participants selecting the correct toy 3–4 times was significantly different by gesture condition, $\chi^2(1, N = 120) = 35.704$, $p < .001$.

We next investigated whether any aspects of the toys or gestures themselves systematically affected performance for the children. It is possible that some gesture types or gesture–toy pairings were simply more transparent than others, or that children have systematic biases to attend to or encode only some information. As mentioned earlier, although we did not intentionally match toys according to whether their gesture encoded manner or path, only one of the four pairs had gestures that included manner. A chi square analysis suggests that the number of correct versus incorrect responses did not vary significantly for that set of gestures as compared to the other three sets, which included only path information, $\chi^2(1, N = 120) = 0.05$, $p = .823$. This suggests that participants' responses were not influenced by whether the gesture encoded manner or path.

The gestures also differed in handshape. Three used a flat palm and five used a pointed index finger. Pointing often serves a deictic function, indicating concrete events and objects in the world, and it is one of the earliest gestures children produce and comprehend (Bates, Bretherton, Shore & McNew, 1983). It is possible that children performed better when the gesture used the flat palm, as this handshape is not as strongly associated with a function. Alternatively, children might perform better when the gestures were points, as they are used to interpreting gestures using this handshape. To explore whether handshape influenced responses, we first compared the number of correct versus incorrect responses for the first toy pairing – the only toy pairing that included both a flat palm gesture and a pointing gesture. There was no difference, $\chi^2(1, N = 30) = 0.159$, $p = .69$. We then compared responses for all of the open palm gestures ($n = 3$) compared to the pointing index finger gestures ($n = 7$). Again, a chi square was not significant, $\chi^2(1, N = 30) = 0.021$, $p = .884$, suggesting that the children were able to interpret gestures including the two handshapes equally well.

Recall that we did attempt to match the toys for 'interestingness'. However, it is possible that some toys still attracted more attention than others, and so we also asked whether performance in the IG condition was systematically different for one of the targets in a toy pairing, collapsing across age groups. This did not appear to be the case; the number of correct responses for the two toys were not significantly different in any set (Set 1: $\chi^2(1, N = 30) = 0.16$, $p = .69$; Set 2: $\chi^2(1, N = 30) = 0.144$, $p = .70$; Set 3: $\chi^2(1, N = 30) = 0.96$, $p = .33$; Set 4: $\chi^2(1, N = 30) = 0.18$, $p = .66$).

The results thus far suggest that both children and adults can use iconic gestures to disambiguate the meaning of novel verbs. However, we cannot conclude from these results alone that participants were actually mapping the novel verb onto the action performed by the toy. Instead, to respond correctly participants only needed to attend to the gesture and directly link the gesture to the toy's action; they could have been responding on the basis of the gesture alone. While that in and of itself is an interesting finding, as it shows they were paying attention to gesture, it does not answer the question of whether gesture can be used as input in word learning. To answer this question, we performed a second study.

Experiment 2

Method

Participants

Ten adults, 3-, and 4-year-olds participated (3-year-olds' mean = 3;6, range: 3;3–3;8; 4-year-olds' mean = 4;2, range: 4;0–4;6).[3] Participants were recruited as in Study 1. No one participated in both studies.

Design and procedure

The task was similar to Experiment 1, but this time we presented a verb and a gesture for both toys in the pair, exposing each participant to eight verbs in total. This required the participant to not only pay attention to the gesture, but also to map the meaning encoded in the gesture onto the correct word. The four additional verbs were *dack*, *gop*, *tam*, and *meek*.

Participants were told that Sam liked both of his toys, but today could only play with one. For example, participants were shown two toys, and then told 'one of Sam's toys lets him (verb + gesture). The other toy lets him (verb + gesture). Sam likes both toys, but today Sam can only play with one of them. Today Sam wants to (verb). Which toy lets Sam go (verb-ing)?' Note that the last two instances of the target verb were not accompanied by a gesture. Note also that, due to counterbalancing, the correct response was not always the action corresponding to the last gesture. Again, the pairing of toy/action and verb was counterbalanced such that the correct response for one half of the participants was the incorrect response for the other half. We also counterbalanced the order of presentation of the sets of toys, such that half the children saw the toy pairs in the reverse order of the other half. And for each individual child we varied the side associated with the correct answer for each toy pair, that is, the correct answer was not consistently on one side or the other.

Results and discussion

Figure 3 shows the percent correct toy choice for Experiment 2, and for comparison, the IG condition from Experiment 1. Again, participants across all ages selected the correct verb at a level significantly above chance (3s: $t(9) = 4.0$, $p = .001$; 4s: $t(9) = 3.0$, $p = .007$; adults: $t(9) = 6.7$, $p < .001$). They were able then, to map the novel verb onto the action performed by the toy. This is all the more impressive given that the relevant gesture was only seen once during this experiment, in contrast to twice in Experiment 1.

Importantly, the overall results of Experiments 1 and 2 were not significantly different from each other (2-factor

[3] Pilot testing suggested the new task was too difficult for 2-year-old children, likely due to the high memory load it requires, and so we did not include them in Study 2.

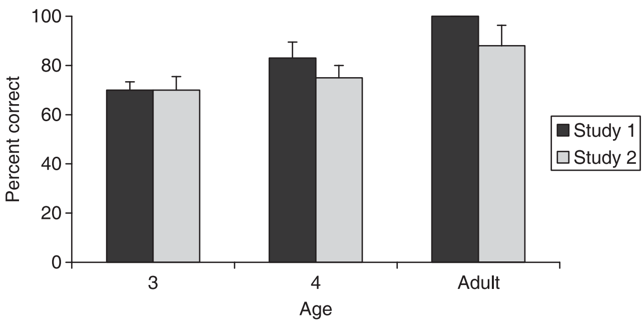

Figure 3 *Percent correct toy choice by age, Experiment 1 vs. Experiment 2.*

ANOVA, age × study: $F(2, 22) = 1.41$, $p = .26$). Thus, although it is possible that participants in Experiment 1 were making their selection based on the gesture alone, the results of Experiment 2 support the possibility that they could have been linking the gesture to the verb in that experiment as well.

As in Experiment 1, we looked to see if the number of participants that got 0–2 correct or 3–4 correct varied by age group. Again, there were slight variations (nine adults, seven 3-year-olds, and seven 4-year-olds performed above chance); however, they were not significant, $\chi^2(1, N = 30) = 0.58$, $p = .45$. We also compared the number of correct responses for each toy in the four sets. As before, there were no significant differences (Set 1: $\chi^2(1, N = 30) = 0.57$, $p = .45$; Set 2: $\chi^2(1, N = 30) = 0.0$, $p = 1.0$; Set 3: $\chi^2(1, N = 30) = 2.4$, $p = .62$; Set 4: $\chi^2(1, N = 30) = 0.96$, $p = .32$). Nor did responses vary systematically according to whether the gesture encoded manner or path. The number of correct versus incorrect responses did not vary significantly for the first set of gestures, which included manner + path, as compared to the other three sets, which included only path, $\chi^2(1, N = 120) = 3.58$, $p = .09$. And we compared responses for all the open palm gestures ($n = 3$) and pointing index finger gestures ($n = 7$). As before, performance did not differ, $\chi^2(1, N = 120) = 0.38$, $p = .53$.

General discussion

Co-speech gestures sometimes contain information corresponding to the message conveyed in speech, but other times information is conveyed uniquely in gesture. Gestures of this latter type provide additional insight into what speakers are thinking, and studies have shown that listeners are sensitive to this information (Goldin-Meadow, 2006). The current findings go beyond previous work by demonstrating that learners are not only sensitive to this information, but that they can use it to interpret unfamiliar parts of speech. Importantly, this ability is not restricted to adults. The children in our studies were at an age when a great deal of language development occurs, particularly vocabulary acquisition. Somewhat to our surprise, there were no real differences between the older and younger children in terms of their ability to use the gestures to interpret the verb. Moreover, the youngest children were able to interpret gestures that they themselves do not yet commonly or frequently produce. Although children produce deictic gestures by 12 months of age (e.g. points at objects) (Bates *et al.*, 1983), iconic gestures tend to come in later, sometime during the second year (Acredolo & Goodwyn, 1988). However, in general young children's iconic gestures are more akin to full body reenactments of an action rather than the hand gestures used by adults (Acredolo & Goodwyn, 1988; McNeill, 1992).

We cannot specify on the basis of the present data how exactly learners might be using gestures. Our assumption is that gestures serve to direct a listener's attention to a given scene (or memory for a scene) wherein an action is contained, and that in general, other cues will be needed to further narrow down the meaning. Verb learning is complicated by many factors, for instance different verbs can refer to the same scene or activity, but take different perspectives on that scene. This ranges from transitive verbs such as *chase* and *flee*, to transitives and intransitives like *push* and *fall* (in a case where someone pushes something causing it to fall). Iconic gestures such as those used in this study will not solve the problem of distinguishing between *chase* and *flee*, although it is possible (indeed even likely) that speakers produce quite different gestures when discussing the pushing versus the falling.

Languages also vary in whether manner or path is encoded in the verb (Talmy, 1985), but here too, gestures seem unhelpful. Some researchers have found that speakers tend to gesture the component not in the verb (McNeill, 2000; McNeill & Duncan, 2000), while others have shown that conflated gestures (including both manner and path) are the most common gesture regardless of language (Özyürek, Kita, Allen, Furman & Brown, 2005). Moreover, syntactic aspects of the sentence may affect what gesturers encode (Özyürek *et al.*, 2005), making the informativeness of the gesture contingent on the syntax. Thus, it would seem that gestures can only get a learner so far, and that other cues, such as knowing the meanings of the nouns (Snedeker & Gleitman, 2004), the syntactic frame in which the verb occurs (Landau & Gleitman, 1985; Naigles, 1990), and acquired sensitivity to language-specific biases (Naigles & Terrazas, 1998) are necessary to help refine the learner's interpretation.

But gestures can get you some traction on the learning problem. In particular, they might help solve the contingency problem mentioned earlier (that verbs tend not to be uttered contemporaneously with the actions they encode), if learners can interpret the iconic gestures that often accompany verbs. In this study we found strong evidence that learners can do this. The current results then join other studies (including Iverson *et al.*, 1999; Özçaliskan & Goldin-Meadow, 2005; and Shatz, 1992) in suggesting that gesture plays a role in language learning.

Acknowledgements

This work was supported by a UC Berkeley Institute of Human Development GSRSF to WG. The authors wish to thank J. Nguyen and N. Tyago for their assistance collecting data, and M. Coppola for comments on an earlier draft.

References

Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development*, **59**, 450–466.

Baldwin, D.A. (2000). Interpersonal understanding fuels knowledge acquisition. *Current Directions in Psychological Science*, **9**, 40–45.

Bates, E., Bretherton, I., Shore, C., & McNew, S. (1983). Names, gestures, and objects: symbolization in infancy and aphasia. In K.E. Nelson (Ed.), *Children's language, Vol. 4* (pp. 59–123). Hillsdale, NJ: Erlbaum.

Bavelas, J. (1994). Gestures as part of speech: methodological implications. *Research on Language and Social Interaction*, **27**, 201–221.

Cassell, J., McNeill, D., & McCullough, K.E. (1998). Speech–gesture mismatches: evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics and Cognition*, **6**, 1–34.

Gentner, D. (1981). Why nouns are learned before verbs. In S. Kuczak (Ed.), *Language development: Vol. 1. Language, thought and culture* (pp. 310–334). Hillsdale, NJ: Erlbaum.

Gleitman, L., & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: the role of syntax in vocabulary acquisition. *Current Directions in Psychological Science*, **1**, 31–37.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Belknap, Harvard University Press.

Goldin-Meadow, S. (2006). Talking and thinking with our hands. *Current Directions in Psychological Science*, **15**, 34–39.

Hadar, U., & Krauss, R. (1999). Iconic gestures: the grammatical categories of lexical affiliates. *Journal of Neurolinguistics*, **12**, 1–12.

Iverson, J., Capirci, O., Longobardi, E., & Caselli, M.C. (1999). Gesturing in mother–child interactions. *Cognitive Development*, **14**, 57–75.

Landau, B., & Gleitman, L. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (2000). Analogic/analytic representations and cross-linguistic differences in thinking for speaking. *Cognitive Linguistics*, **11**, 43–60.

McNeill, D., & Duncan, S. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge: Cambridge University Press.

Morford, M., & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language*, **19**, 559–580.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, **17**, 357–374.

Namy, L., & Waxman, S. (1998). Words and gestures: infants' interpretations of different forms of symbolic reference. *Child Development*, **69**, 295–308.

Namy, L., Campbell, A., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: a U-shaped trajectory in acquisition of arbitrary gestures. *Journal of Cognition and Development*, **5**, 37–57.

O'Neill, D., Topolovec, T., & Stern-Cavalcante, W. (2002). Feeling spongieness: the importance of descriptive gesture in 2- and 3-year-old children's acquisition of adjectives. *Journal of Cognition and Development*, **3**, 243–277.

Özçaliskan, S., & Goldin-Meadow, S. (2005). Do mothers lead their children by the hand? *Journal of Child Language*, **32**, 481–505.

Özyürek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? Insights from crosslinguistic variations and similarities. *Gesture*, **5**, 219–240.

Shatz, M. (1982). On mechanisms of language acquisition: can features of the communicative enviroment account for language development? In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 102–127). New York: Cambridge University Press.

Snedeker, J., & Gleitman, L.R. (2004). Why it is hard to label our concepts. In D.G. Hall & S.R. Waxman (Eds.), *Weaving a lexicon* (pp. 257–293). Cambridge, MA: MIT Press.

Talmy, L. (1985). Semantics and syntax of motion. In T. Shopen (Ed.), *Language typology and syntactic description, Vol. 3, Grammatical categories and the lexicon* (pp. 57–149). Cambridge: Cambridge University Press.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. New York: Cambridge Unviersity Press.

Zukow-Goldring, P., & Ferko, K.R. (1994). An ecological approach to the emergence of the lexicon: socializing attention. In V. John-Steiner, C. Panofsky, & L. Smith (Eds.), *Sociocultural approaches to language and literacy: An interactive perspective* (pp. 170–190). New York: Cambridge University Press.

Received: 26 April 2007
Accepted: 28 November 2007