J. Child Lang. **41** (2014), 462–471. © Cambridge University Press 2013
doi:10.1017/S0305000913000019

BRIEF RESEARCH REPORT

How early do children understand gesture–speech combinations with iconic gestures?*

CARMEN STANFIELD, REBECCA WILLIAMSON, AND ŞEYDA ÖZÇALIŞKAN

Georgia State University

(*Received* 29 *August* 2012 – *Revised* 13 *December* 2012 – *Accepted* 19 *December* 2012 – *First published online* 27 *March* 2013)

ABSTRACT

Children understand gesture + speech combinations in which a deictic gesture adds new information to the accompanying speech by age 1;6 (Morford & Goldin-Meadow, 1992; 'push' + point at ball). This study explores how early children understand gesture + speech combinations in which an iconic gesture conveys additional information not found in the accompanying speech (e.g., 'read' + BOOK gesture). Our analysis of two- to four-year-old children's responses in a gesture + speech comprehension task showed that children grasp the meaning of iconic co-speech gestures by age three and continue to improve their understanding with age. Overall, our study highlights the important role gesture plays in language comprehension as children learn to unpack increasingly complex communications addressed to them at the early ages.

INTRODUCTION

At an early age, children use gesture and speech together to convey meanings that they cannot yet express in speech (Greenfield & Smith, 1976; Goldin-Meadow & Butcher, 2003; Özçalışkan & Goldin-Meadow, 2005a). They, for example, point at a cookie while describing the action to be performed on the cookie (e.g., 'eat' + point at cookie), or point at a book while naming the person to act on the book (e.g., 'mommy' + point at book). In these early combinations the OBJECT (e.g., cookie, book) is almost always indicated by a deictic gesture (Özçalışkan & Goldin-Meadow, 2005a). In contrast,

[*] We thank Georgia State University's Area of Focus Initiative: Research on the Challenges of Acquiring Language & Literacy for providing an undergraduate research fellowship to Carmen Stanfield, and the participating families. Address for correspondence: Şeyda Özçalişkan, Georgia State University, Department of Psychology, PO Box 5010, Atlanta, GA 30302-5010. tel: (404) 413 6282; e-mail: seyda@gsu.edu

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

combinations in which object information is conveyed by an iconic gesture appear much later, typically soon after children begin to show steady increases in their iconic gesture production – roughly around three years of age (Özçalışkan & Goldin-Meadow, 2009, 2011). In this study we ask whether children understand the meaning of gesture + speech combinations – in which an iconic gesture adds object information to the accompanying speech – during the period when they begin to produce similar gesture + speech combinations with iconic gestures in their everyday interactions.

Emergence of iconicity in the child's developing gesture + speech system

Gesture and speech form an integrated system both semantically and temporally in adult speakers; 90% of the gestures produced by adults accompany speech and convey distinct but related information (McNeill, 1992; 2005). In fact, previous work with adult speakers suggests bi-directional interactions between gesture and speech in both comprehension and production, with speech influencing the processing of the accompanying gesture and gesture influencing the processing of the accompanying speech (Kita & Özyürek, 2003; Kelly, Özyürek & Maris, 2010). That is, adult speakers routinely use gesture, particularly iconic gestures, to further extend or to clarify what they convey in speech; similarly they also rely on gesture and speech together to interpret communicative acts directed at them.

Not surprisingly, children learn the gesture + speech system at an early age as well (Goldin-Meadow, 2003). Soon after they begin to produce their first words around the age of one, children use gesture along with speech to convey more complex meanings that they cannot yet express by using speech alone (Goldin-Meadow & Butcher, 2003; Özçalışkan & Goldin-Meadow, 2005a). During this period, deictic gestures, such as pointing at an object to indicate or request that object, constitute the most common gesture type, accounting for roughly 80% of children's overall gesture production (e.g., Iverson, Capirci & Caselli, 1994; Özçalışkan & Goldin-Meadow, 2005a). Thus, between ages one and three, children routinely combine a pointing gesture with a spoken word to convey simple sentence-like meanings ('drink' + point at milk, 'baby' + point at milk), and these early gesture + speech combinations precede and predict the emergence of similar sentence-like meanings in their speech (Iverson & Goldin-Meadow, 2005; Özçalışkan & Goldin-Meadow, 2005b, 2010; see also Greenfield, Lyn & Savage-Rumbaugh, 2008 for a similar report on frequent use of deictic gesture + lexigram combinations in captive bonobos and chimps, our most closely related living primate relatives).

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

In contrast to pointing gestures, the incidence of spontaneous iconic gestures in young children's communicative repertoire is quite rare, typically accounting for less than 5% of the gestures that young children produce (Iverson, Capirci & Caselli, 1994; Nicoladis, Mayberry & Genesee, 1999; Özçalışkan & Goldin-Meadow, 2005b, 2009). The initial significant increase in children's iconic gesture production comes in around age 2;2 (Özçalışkan & Goldin-Meadow, 2011; Özçalışkan, Gentner & Goldin-Meadow, 2013). This spurt in iconic gesture production is followed by increased use of gesture+speech combinations in which the iconic gesture conveys additional meanings not conveyed in speech (e.g., 'Me see it'+ make a V-shape with index and middle fingers to convey RABBIT; 'Talk like this'+place fisted hand next to ear to convey TELEPHONE; Özçalışkan & Goldin-Meadow, 2009). But do we observe a similar increase in children's COMPREHENSION of gesture+speech combinations with iconic gestures around age three as well?

Previous research on children's comprehension of gesture+speech combinations, mostly with deictic gestures, suggests early comprehension. When presented with a gesture+speech combination involving either a deictic ('push'+point at ball) or a conventional gesture ('clock'+GIVE gesture) in a direct requesting context, children aged 1;3 to 2;4 were able to understand the meaning of the gesture+speech combination and offer the requested object to the experimenter (Morford & Goldin-Meadow, 1992). Three- to five-year-old children could even grasp the meaning of an indirect request conveyed in a gesture+speech combination, also with a pointing gesture. For example, upon hearing an adult say, 'It is going to get loud in here' along with a point at the door, children would get up and close the door (Kelly, 2001). Most of this earlier work focused on children's understanding of gesture+speech combinations in which the gesture was either a pointing gesture indicating a concrete referent or a conventional gesture conveying a culturally agreed upon emblematic meaning, thus leaving comprehension of gesture+speech combinations involving iconic gestures relatively unexplored.

We know from previous research that even though children understand the referent of a deictic gesture as early as age 1;0 (Butterworth & Grover, 1988), understanding the iconicity of a gesture is a relatively late achievement, beginning in earnest around age 2;2 (Namy & Waxman, 1998; Namy, 2001; Namy, Campbell & Tomasello, 2004). This late-emerging sensitivity to iconicity might in turn lead to later comprehension of gesture+speech combinations with iconic gestures that convey additional information about an object not found in the accompanying speech. In this study, we explored this possibility by studying two-, three-, and four-year-old children as they participated in a gesture+speech comprehension task depicting familiar everyday objects with iconic gestures and words.

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

In line with previous work that shows late-emerging sensitivity to iconicity, we expected children to show understanding of gesture + speech combinations with iconic gestures at a later age (~three), roughly around the same time they have been shown to use such iconic gestures to produce gesture + speech combinations conveying sentence-like meanings.

METHODS

Participants

Thirty-six children participated in the study, including 12 two-year-olds ($M = 2;2$, range = 2;0–2;4), 12 three-year-olds ($M = 3;5$ months, range = 3;1–3;8), and 12 four-year-olds ($M = 4;4$ months, range = 3;11–4;8), with equal numbers of boys and girls in each age group. The children came from middle- to upper-middle-class families and were predominantly Caucasian (79%) or African American (19%). Data from two additional children were excluded due to experimental error ($N = 1$) or parent prompting ($N = 1$) during data collection.

Procedure for data collection

Each child was tested individually in our laboratory. The experimenter presented the child with six gesture + speech combinations, one at a time. Each combination was composed of an iconic gesture conveying an action characteristic of an object along with a verbal description (e.g., 'I am eating' + move cupped palms towards mouth in parallel as if eating a sandwich). After presenting each gesture + speech combination twice, the experimenter placed a pair of pictures depicting two different objects (e.g., a bowl of cereal vs. a sandwich) on the table, each at an equal distance from the child, and asked the child to choose the picture that matched the description (e.g., 'What did I eat?'). One picture always matched the object depicted in the iconic gesture. For each object pair, half of the children in each age group watched the iconic gesture that matched the action characteristic of the first object (e.g., bowl of cereal) and the other half observed the iconic gesture that matched the action characteristic of the second object (e.g., sandwich). The presentation order of the six picture pairs and the placement of the correct picture (right side vs. left side) were counterbalanced across children. All responses were videotaped.

The picture stimuli consisted of six pairs of laminated cards (11·5 cm × 11·5 cm) depicting different objects. For each pair, the verbal description matched the action associated with both objects (e.g., 'I am eating'). However, the iconic gesture matched only one of the objects in the pair (e.g., scooping movement with a closed fist to convey eating CEREAL vs. moving cupped hand towards mouth with fingers and thumb extended

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

in a U-shape to convey eating SANDWICH). A detailed summary of the spoken and gestural descriptions used for each of the six picture pairs is provided in Table 1.

To ensure that children chose correctly based on their comprehension of the gestures and not because of subtle prompting by the experimenter, we examined the question segment in 25% of the collected data ($N = 54$ test trials, 2 test trials per child). Five independent coders ($M_{age} = 22$), blind to the child's age and research hypothesis, watched the experimenter's behaviors during the question segment in each of the fifty-four selected test trials (i.e., the segment where the experimenter asks 'Which one did I eat?' after placing two pictures in front of the child); the coders were then asked to choose the correct picture (e.g., sandwich vs. cereal) based on what they saw. None of the five coders chose the correct picture significantly above chance ($M_{correct} = 28$ [$SD = 1{\cdot}48$], $M_{incorrect} = 26$ [$SD = 1{\cdot}48$], $t(4) \leqslant {\cdot}81$, $p = {\cdot}15$), and the agreement on picture choice between coders was low across test trials (kappas $\leqslant {\cdot}29$), thus showing that children's responses during testing were not likely to be influenced by subtle prompting by the experimenter.

Procedure for data analysis

We assessed children's responses by identifying which picture the child first indicated, either by pointing, touching, or labeling, from the videos. Each response was scored as correct (1) or incorrect (0), with the total score ranging from 0 to 6 for each child across six test trials. We assessed differences from chance performance with independent *t*-tests, separately for each age group; we also computed differences between age groups by a one-way ANOVA with age as a between-subjects factor. The effect sizes were computed by using eta-squared (hereafter η^2) for the ANOVA comparison and Cohen's *d* (hereafter *d*) for *t*-tests. Reliability on scoring was assessed by re-coding a randomly chosen 25% of the data by a second coder who was blind to the age group and the research hypothesis. Agreement between coders was 100%. Our preliminary analyses showed no significant effect of item, presentation order, or child's sex on the choice score; therefore we collapsed across these factors for subsequent analyses.

RESULTS

Our results showed that children's comprehension of object meanings conveyed through iconic gestures in a gesture + speech combination became evident by age three. As can be seen in Figure 1A, both the three-year-olds and the four-year-olds chose the picture that matched the meaning of the gesture + speech combination significantly above chance; 63% correct for

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

TABLE 1. *Verbal and gestural descriptions used for the stimulus objects pairs*

Item #	Speech	Gesture	Object choice
1	'I'm eating'	Scooping with closed right fist towards mouth	Bowl of cereal
		Moving cupped palms with fingers in U-shape	Sandwich
2	'I'm drinking'	Tilting right closed fist to mouth	Coffee mug
		Tilting right hand in a C-shape to mouth	Glass
3	'I'm playing'	Stacking with claw-shaped hands in alternation	Stack of blocks
		Rotating wrists with closed fists at elbow height	Jump rope
4	'I'm opening'	Rotating extended semi-cupped right hand clockwise while moving it toward body	Door
		Placing touching palms in front of body and opening them	Book
5	'I'm picking'	Moving upward semi-cupped right palm in a half rotation upward	Ball
		Small scooping movement of right hand with index finger and thumb touching	Flower
6	'I'm putting on'	Placing both fisted hands on shoulders and moving them toward chest while rolling shoulders	Jacket
		Gripping the back of the head with left hand; right hand index finger and thumb touching the front of the head, lowering slightly	Hat

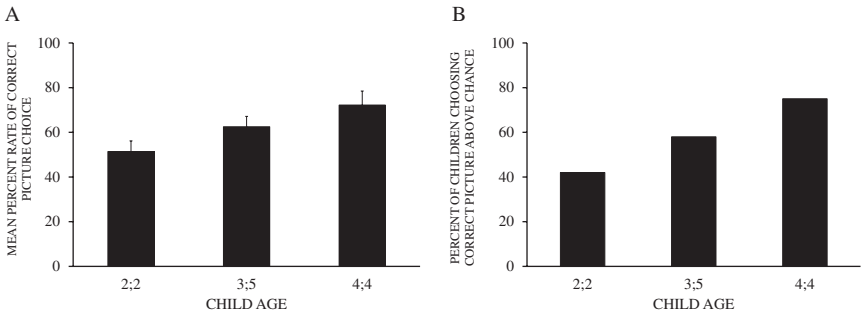


Fig. 1. Mean percentage of children's correct response rate (Panel A) and mean percentage of children providing correct responses above chance levels (Panel B) in the picture-choice task.

the three-year-olds ($t(11) = 2{\cdot}69$, $p = {\cdot}02$, $d = {\cdot}77$) and 72% correct for the four-year-olds ($t(11) = 3{\cdot}55$, $p = {\cdot}01$, $d = 1{\cdot}02$). Two-year-olds, on the other hand, showed no reliable differences from chance performance in choosing the correct picture (51%; $t(11) = {\cdot}29$, $p = {\cdot}77$). The number of children who provided correct responses above chance (i.e., $\geqslant$60% correct response rate) also increased with age: only 42% (5 out of 12) of the two-year-olds chose the correct picture above chance levels, while majority of both the three-year-olds (7/12; 58%) and the four-year-olds (9/12; 75%) were able to identify the correct picture at levels above chance (see Figure 1B).

Children's comprehension of gesture + speech combinations with iconic gestures also improved with age ($F(2, 33) = 3{\cdot}89$, $p = {\cdot}03$, $\eta^2 = {\cdot}19$), from a mean correct response rate of 3·08 ($SD = 1{\cdot}0$) at age two to a mean correct response rate of 4·45 ($SD = 1{\cdot}29$) at age four, revealing a significant difference between ages two and four (Student-Newman-Keuls, $p < {\cdot}05$).

DISCUSSION

In this study we examined whether the comprehension of gesture + speech combinations with iconic gestures emerges around the same time as the production of such gesture + speech combinations in children's communications, and found this to be true. Three-year-olds, but not two-year-olds, were significantly above chance in extracting the meaning of iconic gestures that conveyed unique information not found in the accompanying speech, and this understanding improved with age.

Why do children understand the meaning of gesture + speech combinations with ICONIC GESTURES later than they grasp the meaning of similar gesture + speech combinations with DEICTIC GESTURES? One possible explanation could be the relative cognitive complexity of these two types of

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

gestures. Compared to deictic gestures, the mapping between symbol and referent is less straightforward for iconic gestures. Deictic gestures map onto the world in a direct way by indicating perceptually cohesive entities such as people, objects, or locations. In contrast, iconic gestures select their referents from a set of relational concepts (i.e., associated actions and attributes) and, as such, might impose additional cognitive challenges for young children (Gentner, 2006; Özçalışkan *et al.*, 2013).

A second possible explanation for the late comprehension of gesture + speech combinations with iconic gestures might be the lack of familiarity with such gestures at the early ages. Previous research has shown that even though children can associate iconic gestures with objects at age 1;6, it is not until age 2;2 that they truly understand the iconic relation between gesture and referent (Namy & Waxman, 1998; Namy, 2001; Namy, Campbell & Tomasello, 2004). In fact, deictic gestures typically precede children's first words (Iverson & Goldin-Meadow, 2005), whereas iconic gestures appear several months after children produce their first words (Özçalışkan *et al.*, 2013). Moreover, the incidence of iconic gestures in children's immediate environment is very scarce; parents of young children rarely use iconic gestures when communicating with their young children, and instead rely predominantly on pointing gestures (Iverson *et al.*, 1994; Iverson, Capirci, Longobardi & Caselli, 1999; Özçalışkan & Goldin-Meadow, 2005b, 2011). As such, comprehension of gesture + speech combinations with iconic gestures might be particularly challenging for young children who have not yet themselves produced or observed others produce a substantial amount of iconic gestures.

In fact, even the four-year-olds in our study were performing below ceiling levels, suggesting a more extended developmental trajectory for children's comprehension of gesture + speech combinations with iconic gestures. Previous work on gestures that accompany children's early narrative productions shows that children go through a steep increase in their PRODUCTION of iconic co-speech gestures – particularly the ones that convey story events from the story character's viewpoint – around age five (Stites & Özçalışkan, 2012; see also McNeill, 1992: 295–328 for a related discussion). A similar trajectory might be true for the COMPREHENSION of gesture + speech combinations with iconic gestures as well, with initial significant improvements in comprehension becoming evident around age three – at the same time as the initial spurt in iconic gesture production (Özçalışkan *et al.*, 2013), which is then followed by a more full-fledged understanding of such combinations around age five.

Previous work shows clear evidence of early comprehension abilities of gesture + speech combinations with deictic gestures (e.g., 'push' + point to ball; Morford & Goldin-Meadow, 1992). Our study extends this finding by showing that children also develop the ability to understand

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

gesture + speech combinations with ICONIC GESTURES, but at a slightly later age, during the period when they also begin to produce such combinations in their communications. Overall, our results suggest that the complexity of a gesture (ICONIC vs. DEICTIC) as well as its relative frequency in the input might serve as important contributors to children's developing ability to unpack multimodal communications addressed to them at the early ages.

REFERENCES

Butterworth, G. E. & Grover, L. (1988). The origins of referential communication in human infancy. In L. Weiskrantz (ed.), *Thought without language*, 5–24. Oxford: Clarendon.

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek & R. Michnick Golinkoff (eds.), *Action meets word: how children learn verbs*, 544–64. New York: Oxford University Press.

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

Goldin-Meadow, S. & Butcher, C. (2003). Pointing toward two-word speech in young children. In S. Kita (ed.), *Pointing: where language, culture, and cognition meet*, 85–107. New Jersey: Erlbaum Associates.

Greenfield, P. & Smith, J. (1976). *The structure of communication in early language development*. New York: Academic Press.

Greenfield, P. M., Lyn, H. & Savage-Rumbaugh, S. (2008). Protolanguage in ontogeny and phylogeny: combining deixis and representation. *Interaction Studies* **9**, 34–50.

Iverson, J. M., Capirci, O. & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development* **9**, 23–43.

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

Iverson, J. M. & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science* **16**, 367–71.

Kelly, S. D. (2001). Broadening the units of analysis in communication: speech and nonverbal behaviors in pragmatic comprehension. *Journal of Child Language* **28**, 325–49.

Kelly, S. D., Özyürek, A. & Maris, E. (2010). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychological Science* **21**, 260–67.

Kita, S. & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* **48**, 16–32.

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

McNeill, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.

Morford, M. & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language* **9**, 559–80.

Namy, L. L. (2001). What's in a name when it isn't a word? 17-month-olds' mapping of nonverbal symbols to object categories. *Infancy* **2**, 73–86.

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

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

Nicoladis, E., Mayberry, R. I. & Genesee, F. (1999). Gesture and early bilingual development. *Developmental Psychology* **35**, 514–26.

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press

Özçalışkan, Ş., Gentner, D. & Goldin-Meadow, S. (2013). Do iconic gestures pave the way for children's early verbs? *Applied Psycholinguistics*. Available on CJO. doi:10.1017/S0142716412000720.
Özçalışkan, Ş. & Goldin-Meadow, S. (2005a). Gesture is at the cutting edge of early language development. *Cognition* **96**, 101–13.
Özçalışkan, Ş. & Goldin-Meadow, S. (2005b). Do parents lead their children by the hand? *Journal of Child Language* **32**, 481–505.
Özçalışkan, Ş. & Goldin-Meadow, S. (2009). When gesture–speech combinations do and do not index linguistic change. *Language and Cognitive Processes* **28**, 190–217.
Özçalışkan, Ş. & Goldin-Meadow, S. (2010). Sex differences in language first appear in gesture. *Developmental Science* **13**, 752–60.
Özçalışkan, Ş. & Goldin-Meadow, S. (2011). Is there an iconic gesture spurt at 26 months? In Gale Stam & Mika Ishino (eds.), *Integrating gestures: the interdisciplinary nature of gesture*, 163–74. Amsterdam: John Benjamins.
Stites, L. J. & Özçalişkan, Ş. (2012). Narrative perspective appears first in gesture. Poster presented at the 37th *Boston University Conference on Language Development*, Boston, MA.

https://doi.org/10.1017/S0305000913000019 Published online by Cambridge University Press