The relationship between character viewpoint gesture and narrative structure in children*

FEY PARRILL

BRITTANY LAVANTY

AUSTIN BENNETT

ALAYNA KLCO

Department of Cognitive Science, Case Western Reserve University

AND

OZLEM ECE DEMIR-LIRA

Department of Psychology, University of Chicago

ABSTRACT

When children tell stories, they gesture; their gestures can predict how their narrative abilities will progress. Five-year-olds who gestured from the point of view of a character (CVPT gesture) when telling stories produced better-structured narratives at later ages (Demir, Levine, & Goldin-Meadow, 2014). But does gesture just predict narrative structure, or can asking children to gesture in a particular way change their narratives? To explore this question, we instructed children to produce CVPT gestures and measured their narrative structure. Forty-four kindergarteners were asked to tell stories after being trained to produce CVPT gestures, gestures from an observer's viewpoint (OVPT gestures), or after no instruction in gesture. Gestures were coded as CVPT or OVPT, and stories were scored for narrative structure. Children trained to produce CVPT gestures produced more of these gestures, and also had higher narrative structure scores compared to those who received the OVPT training. Children returned for a follow-up session one week later and narrated the stories again.

The training received in the first session did not impact narrative structure or recall for the events of the stories. Overall, these results suggest a brief gestural intervention has the potential to enhance narrative structure. Due to the fact that stronger narrative abilities have been correlated with greater success in developing writing and reading skills at later ages, this research has important implications for literacy and education.

1. Introduction

Children's gestures are very informative about their language learning (Goldin-Meadow, 2005; Iverson & Goldin-Meadow, 2005; Özçalışkan & Goldin-Meadow, 2009). This is true not just for the acquisition of words and syntax, but also appears to be true of children's ability to tell a well-structured story (that is, a story structured around sequences of goals, attempts, and outcomes: Demir, Levine, & Goldin-Meadow, 2014). Demir and colleagues (2014) found that, when telling a story, five-year-old children who spontaneously gestured from the point of view of a character produced more cohesive and coherent narratives at later ages, or had better NARRATIVE STRUCTURE. In their study, children were asked to retell a story at four different ages, beginning at age five and ending at age eight. The authors explored which factors best predicted narrative structure at later ages. They found that gestures from the point of view of a character predicted well-structured narratives. Demir and colleagues' work suggests an intriguing possibility: If children who spontaneously gesture from the point of view of a character tell better stories later on, does asking children to produce such gestures improve their storytelling? In the present study, we ask whether children who are instructed to gesture from the point of view of a character when telling stories display better narrative structure in those stories. Before describing the study, we briefly discuss the relevance of narrative structure to language and literacy development in later years, summarize relevant research on narrative structure and narrative development, and discuss the relationship between gesture, perspective taking, and cognition.

Narrative abilities have important implications for children's literacy, as they are considered a bridge between earlier developing oral language skills and later developing reading skills. Individual differences in narrative skill relate to children's later reading success (Feagans & Short, 1984; Gillam & Johnston, 1992; Griffin, Hemphill, Camp, & Palmer Wolf, 2004). Given the importance of narrative for later reading success, promoting narrative skills is a major focus in research on early literacy. While results thus far have been mixed, it does seem possible to design interventions that promote narrative development in young children (Petersen, 2011; Peterson, Jesso, & McCabe, 1999).

Children start referring to past events and mentioning future events at two years of age (McCabe & Peterson, 1991). By ages five to six, children begin to produce their own independent narratives organized around goals and attempts of characters, and they reliably and frequently refer to the main components of narrative structure (initial orientation, complication, and resolution: Berman & Slobin, 1994). These ages constitute significant transitional years where children's narrative skills vary greatly (Berman & Slobin, 1994; Colletta, 2009; Reilly, 1992). Children's narrative structure continues to develop during the school years, extending into adolescence (Applebee, 1978; Berman & Slobin, 1994; Warden, 1976).

Narratives are structured at both the micro- and macro-levels (Ninio & Snow, 1996). COHESION (organization at the micro-level) refers to the linguistic, local relations that tie the span of idea units in the narrative together and create a text (Halliday & Hasan, 1979; Karmiloff-Smith, 1979). For example, anaphoric pronouns like *it* create cohesion by tying two nominal elements together. The focus of the current paper is on narrative structure at the macro-level. COHERENCE (organization at the macro-level) refers to elements that give narrative content a schematic organization and tie the different parts of the narrative to each other in a meaningful way (Bamberg & Marchman, 1990; Berman & Slobin, 1994). For instance, we might use general world knowledge, context, and inference to fill in gaps when interpreting a narrative, as well as using the linguistic features that give rise to cohesion (Brown & Yule, 1983). Of course, micro- and macro-level structure are not separable, as many of the elements that create cohesion also create coherence (Dancygier, 2012). Macro-level narrative structure has been defined in the literature in different ways; commonalities among different definitions are that well-structured narratives are organized around the goals and attempts of the story's characters, and that a particularly salient feature of well-structured narratives is that they conform to a hierarchical story schema with various episodes. Each episode consists of an initiating event that leads to the creation of a goal by the character, an attempt to achieve the goal, failure or success in achieving the goal, and reactions of the character to the consequence. The components that make up episodes are temporally or causally related (Berman & Slobin, 1994; Labov & Waletzky, 1997; Stein & Glenn, 1979).

The literature on narrative development has focused primarily on children's speech, but children often produce gestures along with their narratives (Alibali, Evans, Hostetter, Ryan, & Mainela-Arnold, 2009; Colletta et al., 2015; McNeill, 1992). These gestures have the potential to serve as a unique window into children's narrative development (Cassell & McNeill, 1991). Berman and Slobin (1994) and Reilly (1992) suggest that young children might express their early representations through extralinguistic means (such as gesture) before they become adept at telling a story. Being asked to gesture in particular

410

ways has been shown to promote language skills for earlier-developing aspects of language, such as vocabulary (LeBarton, Raudenbush, & Goldin-Meadow, 2015). Might being instructed to gesture from the point of view of a character promote the creation of narrative structure?

As noted already, another aspect of gesture that has important implications for both narrative structure and cognition is viewpoint. Viewpoint in gesture has been shown to correlate with particular discourse contexts (Debreslioska, Özyürek, Gullberg, & Perniss, 2013; Parrill, 2010). Character viewpoint (hereafter CVPT) gestures occur at particularly central moments in a discourse, so asking participants to produce those kinds of gestures might encourage the mention of central events. More specifically, Demir and colleagues hypothesized that CVPT gesture indicates that a child is imagining the event from the perspective of the character. In their 2014 study, children's narratives and gestures were studied at ages five through eight. While children who produced CVPT gestures at age five did not have better narrative structure at that time, they did produce better-structured narratives at ages six, seven, and eight. Children's OVPT (observer viewpoint) gestures did not predict narrative structure at later ages. Thus, Demir and colleagues argue that there is something special about CVPT gesture. They speculate that imagining the event from the perspective of the character results in greater focus on that character's goals and attempts.

Support for the claim that taking a character (or actor) viewpoint changes encoding can also be found in the work of Brunyé and colleagues (Brunyé, Ditman, Mahoney, Augustyn, & Taylor, 2009; Brunyé, Ditman, Mahoney, & Taylor, 2011, Ditman, Brunyé, Mahoney, & Taylor, 2010). In several studies, these researchers found that people can be prompted to take an actor's viewpoint on an event, and that doing so changes encoding in ways that are parallel to experiencing the event as an actor. After reading narratives containing second person pronouns (*you are …*), participants were faster to recognize pictures of actions depicted from the point of view of an actor, as compared to pictures showing an observer's view of the action (Brunyé et al., 2009). Being presented with an action from the point of view of an actor also led to better memory for that action (Ditman et al., 2010). Further, reading about negative events from a *you* perspective (leading to a first person imagining of the event) also changed participants' mood (Brunyé et al., 2011). In these studies, linguistic cues were used to encourage first person encoding of events, and this appeared to be effective in changing encoding. These studies focus on speech, but other work has found re-enacting gesture from CVPT to benefit recall (Wesson & Salmon, 2001). This finding has to be disentangled from a general benefit of being asked to gesture on recall (Cook, Duffy, & Fenn, 2013; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014, Stevanoni & Salmon, 2005). However, if CVPT gesture does

411

boost recall, this might help to explain why it predicted better narrative structure at later ages. That is, children instructed to gesture from the point of view of a character might imagine the cartoon from an agent's point of view, and therefore have better recall for details of the stories, as well as better narrative structure.

The inter-relationships hypothesized by Demir and colleagues (2014) can be schematized as in Figure 1. For whatever reason, possibly better perspective-taking ability, some children imagine certain events from a first person point of view. This results in the production of CVPT gesture, and also in better encoding of the event (better focus on goals and actions). Only after the maturation of narrative ability (linguistic competence, short-term memory capacity) does this difference result in better-structured narratives.

However, there is no direct evidence that CVPT gesture can be used to infer a first person point of view. Perhaps the closest is a study by Parrill and Stec (2018). In this study, participants read stories containing either second person pronouns (*you are …*) or third person pronouns (*she is …*), and then narrated the stories. Participants who read the second person versions produced first person speech. They did not produce more CVPT gestures compared to participants who received third person descriptions. Because being prompted to produce first person speech did not change gesture, the question of what exactly CVPT gesture means about a person's conceptualization remains open.

The current study attempts to tease apart some of the relationships schematized in Figure 1. We asked children to produce CVPT gesture in order to see if this instruction impacts narrative structure scores. If asking children to produce CVPT gesture causes them to take a first person point of view, this instruction could result in better encoding for the events. Better encoding
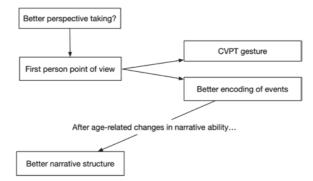

Fig. 1. Possible relationships among first person point of view, CVPT gesture, encoding, and narrative structure.

412

could in turn lead to better narrative structure and better recall. This account (schematized in Figure 2) argues that being prompted to gesture from a particular point of view changes how children conceptualize the event in a way that also improves narrative structure and recall.

If children who are asked to produce more CVPT gestures do not have better narrative structure and recall, there are several possible explanations. It may be that simply producing CVPT gesture does not result in a first person point of view. It may be that producing CVPT gestures does result in a first person point of view, but the first person point of view does not lead to better encoding. It may also be that further development of narrative ability is necessary, and that gestural intervention cannot impact children of this age.

In summary, this study attempts to 'reverse the arrow' between first person point of view and CVPT gesture that is shown in Figure 1. Addressing the relationship between character viewpoint gesture and narrative ability can shed light on the conceptual processes that underlie them both. In addition, improving children's narrative skills is of educational importance, given the strong relationships between early narrative skills and later reading success (Feagans & Short, 1984; Gillam & Johnston, 1992; Griffin et al., 2004). The improvement of narrative structure thus has important implications for developing literacy skills later on. Our main research questions are: Compared to other kinds of gestures or a no instruction control, does asking children to gesture from the point of view of a character (1) improve their narrative structure immediately, (2) improve their narrative structure after a one-week delay, and (3) lead to better recall?

2. Method

2.1. PARTICIPANTS

Forty-four kindergarten students from a private school in the northeastern US participated in the study (23 females, mean age at the time of the session six years, one month, range 5;3–6;9). Parents signed a consent document
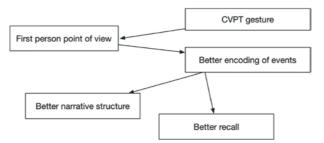

Fig. 2. Hypothesized effects of CVPT gesture on encoding, narrative structure, and recall.

413

indicating their willingness to have the child participate, and also filled out a demographic questionnaire. Demographic data are included in the 'Appendix'. The child was also asked to give oral consent at the start of the study. No compensation was provided to parent or child for participating. Each child participated in two sessions, a narrative session and a recall session one week later. Both sessions were videotaped.

2.2. MATERIALS

Children watched three training cartoon video clips and four experimental cartoon video clips. Clips came from cartoons, a common way to elicit gesture data for both children and adults (see, e.g., Demir et al., 2014; McNeill, 1992). While they have drawbacks (e.g., anthropomorphized animals might not be a good data source for our understanding of human mental representations), they also have significant strengths. For instance, they involve actions and movements that can be easily schematized in gesture, and participants tend to find them engaging. We used cartoon video clips for this study because they tend to elicit iconic gesture (gesture that closely correlates to the semantic content of speech), and specifically gestures from the point of view of a character (CVPT), and an observer (OVPT). As noted by Parrill (2010), not all events can be gestured from both a character and an observer point of view. For example, a person holding a newspaper is very likely to evoke a CVPT gesture showing the hands holding an object, and is unlikely to be gestured from OVPT. A gesture showing a complex trajectory (e.g., a character running down a mountain) is more likely to evoke an OVPT gesture. Some events have the potential to evoke both types of gesture. A character swinging through space on a rope can be gestured by showing the character's hands holding the rope and moving (CPVT), or by showing the path the character took by tracing it (OVPT). For both the training and experimental session we used video clips that contained multiple events that could be gestured from either point of view. Because we were exploring narrative structure, we also made sure that all videos used in the experimental trials contained multiple goal–attempt–outcome sequences.

2.3. PROCEDURE: TRAINING

Children participated individually. When they entered the experiment room, the experimenter introduced a 'story-telling circle' (see Figure 3) by saying, *You probably noticed there's this circle on the floor. This is a special spot we're going to use to tell stories. Let me show you.* She then stepped into the circle and told a brief story during which she produced unscripted gestures (no effort was made to keep them the same every time for this introduction, though the

414

Fig. 3. Experiment set-up.

story itself was the same). The story-telling circle helped ensure that children stood in the same place for each trial, while having the experimenter model story-telling helped them feel more comfortable and helped them get a sense of what was going to happen in the study. The experimenter then asked for oral consent from the child by saying, *I brought some of my favorite cartoons to show you. Do you want to watch them? Do you want to ask any questions before we watch them?*

Following consent, the child participated in three training trials. The training and experimental procedure used a between-participants design, so that each child was in one condition throughout. During the three training trials, the experimenter and child watched a five-second video clip together; the experimenter described the clip in a way that differed across conditions, then the child was prompted to describe the clip according to specific instructions that differed by condition. The training video clips were segments from videos not used in the experimental portion of the study. Differences across conditions are as follows:

1. Character viewpoint condition (tell, show, and pretend). After watching the clip, the experimenter said *Now I'm going to tell you and show you what happened and I'm going to pretend I'm the* [character, e.g., *mouse, cat*]. The experimenter acted out two specific events from the clip, each paired with a specific character viewpoint gesture. An example is provided in Table 1. The experimenter then prompted the child to produce a description by saying: *Now can you tell me and show me what happened and pretend you're the* [character]?

2. Observer viewpoint condition (tell and show). After watching the clip, the experimenter said, *Now I'm going to tell you and show you*

415

TABLE 1. *Narrative structure scoring* (*Demir et al.,* 2014)

Category	Description
Descriptive sequence	Are story events temporally organized, i.e., the events follow one another in time. If no (story does not include a temporal structure and only contains the physical and personality characteristics of an animate protagonist), categorized as a descriptive sequence.
Action sequence	If story does have temporal structure, does it have causal structure (i.e., one event causing the following event or events)? If no, categorized as action sequence.
Reactive sequence	If story does have causal structure, does it have a goal (expressed as character 'wants to', 'tries to', does something 'in order to', or 'because of', or 'so')? If no, categorized as reactive sequence.
Incomplete goal-based sequence	If the story has temporal structure, causal structure, and goals and/or actions the protagonist performed to achieve the goals, does it have outcome (i.e., whether the protagonist reached the goal or not)? If no, categorized as incomplete goal-based story.
Complete goal-based story with one episode	If the story does have an outcome, does it have more than one goal–attempt–outcome sequence? If no, categorized as complete goal-based story with one episode.
Complete goal-based story with multiple episodes	If story has multiple goal–attempt–outcome sequences, categorized as complete goal-based story with multiple episodes.

what happened. The experimenter acted out two specific events from the clip, each paired with an observer viewpoint gesture (see Table 1). The experimenter prompted the child to produce a description by saying: *Now can you tell me and show me what happened?*

3. Control condition (tell). After watching the clip, the experimenter said, *Now I'm going to tell you what happened*. The experimenter repeated the verbal descriptions from the same two specific events from the clip, this time with no gesture. The experimenter prompted the child to produce a description by saying: *Now can you tell me what happened?*

In summary, the children were first trained with three training trials. In the character viewpoint condition, children were prompted to gesture (*show me*), and to take on the role of the character (*pretend you are the …*), and character viewpoint gestures were modeled for them. In the observer viewpoint condition, children were prompted to gesture (*show me*), but they were not prompted to take on the role of the character, and observer viewpoint gestures were modeled for them. In the control condition, children were not prompted to gesture (*tell me …*), though they could still choose to do so. The gestures performed by the experimenter were a character viewpoint variant and an observer viewpoint variant of the same event across the gesture conditions,

and were always performed the same way within conditions. An example is given in the 'Appendix'.

In the character and observer viewpoint conditions, if the child produced a gesture of the correct type (character in the character viewpoint condition, observer in the observer viewpoint condition), the experimenter said *That was great!* and moved on to the next training trial. If the child produced gesture but no speech, the experimenter prompted her or him with *I like how you showed me, but can you tell me?* If the child produced speech but did not produce the correct type of gesture in one of the gesture conditions, the experimenter prompted the child again, using the respective prompts for each condition, as provided in the 'Appendix'. After the child's second attempt, the experimenter moved on to the next training trial regardless of whether or not the child produced the correct kind of gesture. It should be noted that the narrations children produced during training were not part of the analysis.

2.4. PROCEDURE: EXPERIMENTAL TASK

After the child completed the three training trials, the experimenter moved on to the experimental task, again using a between-participants design in which the child received prompts that were the same as during training. In this task, the child watched four one-minute video clips. While these short video clips did not allow for very extended narratives, each contained multiple goal–attempt–outcome sequences, and were appropriate for use with young children. After watching each clip, the children were prompted to describe the clip to the experimenter. The experimenter did not produce any gestures during the experimental task. Prompts began by asking the child if he or she noticed the main character: *There was an* [animal] *in the story! Did you see the* [animal]*?* The experimenter clarified if necessary. Following this, the prompt used varied by condition.

1. Character viewpoint condition description prompt. *Can you tell me and show me what happened and pretend you're the* [animal]*?*
2. Observer viewpoint condition description prompt. *Can you tell me and show me what happened?*
3. Control condition description prompt. *Can you tell me what happened?*

If the child produced a very incomplete narration (e.g., *there was a dog*) the experimenter prompted once with *and then what happened?* The number of children in the three conditions were as follows: character viewpoint: 15, observer viewpoint: 15, control: 14.

After each description, children were asked a set of comprehension questions (shown in the 'Appendix'). The purpose of these questions was to assess whether the children's basic understanding of the events in the stories

417

was the same across conditions. The prompt for the comprehension questions was always: *That was a great story! I'm going to ask you some questions about the* [animal]. After completing all four trials (description followed by comprehension questions), the experimenter told the child: *Okay, I'm going to see you again in a week and we're going to do some other stuff, but that's it for today! Thank you so much for watching these cartoons with me!*

The training and experimental trials were presented using Superlab (Cedrus Corporation) experiment software, presented on a laptop. The children stood to watch the videos while the experimenter sat next to them, they then moved to the story-telling circle while the experimenter sat next to it, as shown in Figure 3.

2.5. PROCEDURE: RECALL SESSION

In the recall session, each child did the following:

1. A FREE RECALL SESSION. The child was prompted to remember the four experimental videos with the following prompt: *Last time I saw you we watched a cartoon about a* [animal]. *Tell me what happened in that cartoon and tell me everything you can remember.*

2. A CUED RECALL SESSION. The child was prompted to recall the four experimental videos with the following prompt: *Here's a picture from one of the cartoons we watched.* [The child was shown a still image of the main character in the story engaged in an event from the story.] *Tell me again: What happened in the cartoon, and tell me everything you can remember.*

3. A SYNTAX COMPREHENSION TASK. The experimenter administered a picture-matching syntax comprehension task (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). Details and a sample item can be found in the 'Appendix'. This task provides a measure of spoken language comprehension.

4. A SPATIAL MEASURE. The experimenter administered the short form (16 trials) of the Thurstone Primary Mental Abilities Spatial Relations Subtest (Thurstone, 1974), a basic measure of spatial ability. Details and a sample item can be found in the 'Appendix'. We explored spatial skill because of suggestions that high spatial skill is associated with higher gesture rate (especially when combined with low ability to quickly produce words – phonemic fluency; Hostetter & Alibali, 2007). As with the narrative session, the recall still images and the syntax comprehension pictures were presented using Superlab experiment software, presented on a laptop. The story-telling circle was also used for the recall tasks (see Figure 3). The spatial measure was done on paper.

418

2.6.1. *Speech and gesture*

Speech and gesture coding was carried out in ELAN (Crasborn & Sloetjes, 2008; ELAN is an annotation software program created by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, the Netherlands; see <http://tla.mpi.nl/tools/tla-tools/elan/>). Speech was transcribed and divided into clauses. We defined a clause as an expression of a single situation, typically containing a subject (which may have been unexpressed syntactically) and a predicate. We averaged over the stories to obtain mean number of clauses for each child.

All gestures produced by the children were transcribed and sorted into the following categories (after McNeill, 1992): CONCRETE ICONIC (features of the hand or body action map onto features of the stimulus event), or OTHER. 'Other' included gestures that depict abstract content such as ideas or emotions (often called METAPHORIC gestures), DEICTIC (gestures that point or locate objects in space), and BEATS (rhythmic gestures with little additional semantic content). Because of our specific research questions, we did not include other gestures in our analysis.

Concrete iconic gestures were then coded for viewpoint. Gestures in which children took on the role of the character, using their bodies as the character's body, were identified as character viewpoint gestures (CVPT), as shown in Figure 4a. Gestures in which children used the hand or arm to reflect the character as a whole, as though showing the scene as an observer, were identified as observer viewpoint gestures (OVPT), as shown in Figure 4b.

We averaged over the stories to obtain mean number of gestures, and mean proportion of CVPT and OVPT for each child. We also calculated gesture rate for each child as the mean number of gestures divided by the mean number of clauses.

2.6.2. *Narrative structure score (NSS)*

Using an ELAN output transcript of each child's speech, narrative structure was coded using the process described by Demir and colleagues (2014, p. 667). Their scheme comes from Stein and Glenn (1979). This scheme is shown in Table 1.

It should be noted that there are multiple ways to measure narrative complexity or strength. For example, some researchers (Nicoladis, Marentette, & Navarro, 2016) focus on the use of clausal connectors (e.g., *and*, *or*, *so*, *when*) as a measure of narrative complexity. Others focus on number of clauses, uses of connectors, and also on anaphoric expressions (Colletta, 2009). We used a narrative structure scoring system that is organized around goal, attempt, and outcome sequences because we wanted

419

A. He goes up a flagpole   B. He climbed up the flagpole

Fig. 4. Character and observer viewpoint gestures.

to extend the findings of the previous study that made use of this scheme (Demir et al., 2014), not because we regard it as the most definitive scheme. However, it also important to note that there are common elements to the schemes used by researchers studying multimodal language, particularly the use of clausal connectors (*because*, *when*) that indicate temporal and causal relationships.

Stories received a numerical score based on which category they were placed in. A story categorized as a descriptive sequence received a 1, a story categorized as an action sequence received a 2, and so on. The highest score a story could receive was therefore a 5. However, despite the fact that videos contained multiple goal sequences, in our case no stories were scored as 5. Because children narrated four stories, we averaged their narrative structure scores across the stories to yield a mean narrative structure score. Examples of each category of story are presented in the 'Appendix'.

2.6.3. *Perspective in speech*

As in Demir and colleagues (2014), we coded direct and indirect quotations. We also identified expressions of cognition and emotion in children's speech, both of which indicate that the child is taking the point of view of the character. We used the scheme of Langdon, Michie, Ward, McConaghy, Catts, and Coltheart (1997) and identified any lexical terms describing perception (e.g., *see*), desire (e.g., *want*), emotion (e.g., *feel*), and cognition (e.g., *think*) (Langdon et al., 1997, p. 176).

This coding was done for each clause, and a single clause could have more than one of these categories present. A child's count was then totaled across these categories (direct quotes, indirect quotes, perception/cognition verbs),

420

and averaged over stories to yield a measure indicating mean number of instances of first person perspective in speech.

2.6.4. *Comprehension question scoring*
Each question was scored as correct or incorrect. This score was then divided by number of questions for proportion correct.

2.7. CODING: RECALL SESSION
Four children (two control, two CVPT) were not used in the recall session because video data were corrupted or lost after a hard disk failure. Speech, gesture, narrative structure score, and first person speech for the recall session were coded as described above. To measure recall, we used the method of Stevanoni and Salmon (2005), which was itself adapted from Jones and Pipe (2002). Children got a point for every core action and core object mentioned. For example, in describing a scene where the coyote puts on a bib, if the child says *the coyote puts on a bib*, he or she would get a point for core action (*put on*) and a point for core object (*bib*). We also counted the number of distortions or intrusions (child incorrectly mentions a core action or an object). For example, if in describing a scene where the coyote picks up a fork and knife the child says the coyote picked up a spoon, this would be scored as a distortion of the core objects knife/fork. If the child said *the coyote threw a fork*, this would be scored as a distortion of the core action *pick up*. We did not count a description as an error if the child simply didn't know how to describe it (e.g., the roadrunner was described as *bird*). For the cued recall scoring, the child got a point only for any new core action or core object mentioned, or for new distortions/intrusions. Because we have no hypotheses about cued versus free recall, we collapsed over these parts of the session to create total recall measures for actions/objects and distortions/intrusions. Finally, to create a single measure of recall, we subtracted distortions/intrusions from actions/objects recalled.

2.7.1. *Syntax comprehension scoring and PMA Spatial Relations Subtest scoring*
For both the syntax comprehension and the spatial relations test, children received a score indicating percent correct (number correct divided by number of questions).

2.8. RELIABILITY
We randomly selected seven children from the completed dataset (15% of the data) and a second coder coded both narrative and recall data to establish

reliability of coding for gesture type (97% agreement), narrative structure (100% agreement), perspective in speech (96% agreement), and correct recall (94% agreement).

3. Results

We present results for the narrative session first, followed by results for the recall session. We discuss our reasons for using particular statistical procedures and other statistical detail in the 'Appendix'. Data are available via the Open Science Framework <https://osf.io/6ds5k/>. A table with descriptive statistics for all measures can be found in the 'Appendix'.

3.1. NARRATIVE SESSION

Our key research questions are whether training impacted children's gestures, and whether training impacted narrative structure scores. That is, was there an effect of the training condition on gesture and on narrative structure score? Figure 5a shows the mean proportion CVPT gesture by condition (recall that proportion CVPT + proportion OVPT = 100%). Figure 5b shows mean narrative structure scores by condition.

Because we are interested in the effect of the training condition on two outcome variables, we used a one factor (condition) MANOVA with proportion CVPT and mean narrative structure score as outcome variables to analyze our results. We found a statistically significant effect of condition ($F(4,82) = 5.8$, $p = .001$, $\eta_p^2 = .22$). Tests of between subjects effects showed a significant difference across conditions for both proportion CVPT ($F(2,41) = 12.98$,
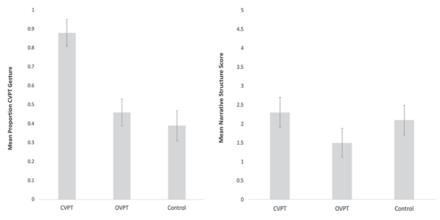
Fig. 5. Mean proportion CVPT gesture (left) and mean narrative structure score (right) by condition; error bars show standard error of the mean.

422

p = .001, η_p^2 = .39) and narrative structure score ($F(2,41)$ = 3.52, p = .001, η_p^2 = .15). Planned post-hoc comparisons were conducted for both dependent variables using Scheffé's test (because of unequal group sizes). For CVPT gesture production, the CVPT condition differed significantly from both the OVPT (mean difference = 0.42, 95% CI [0.15, 0.68]) and the control (mean difference = 0.42, 95% CI [0.22, 0.76], while OVPT and control did not differ from each other: mean difference = 0.06, 95% CI [–0.20, 0.34]. For narrative structure score, the CVPT condition differed significantly from the OVPT (mean difference = 0.87, 95% CI [0.03, 1.7]), but not the control (mean difference = 0.50, 95% CI [–0.35, 1.3]). OVPT and control did not differ from each other: mean difference = 0.36, 95% CI [–0.49, 1.2].

These initial analyses suggest that being asked to gesture from the point of view of the character did have an effect on CVPT gesture and narrative structure scores. However, it is also possible that differences in overall gesture rate explain the patterns we see. To address this possibility, we carried out a one factor ANOVA (condition) with gesture rate (number of gestures / number of clauses) as the dependent variable. There was a significant difference in gesture rate across conditions ($F(2,41)$ = 6.97, p = .002, η_p^2 = .25). Planned post-hoc comparisons showed that the CVPT condition children had a higher gesture rate than control (mean difference = 0.23, 95% CI [0.03, 0.44]), but not OVPT (mean difference = –0.04, 95% CI [–0.25, 0.16]). OVPT children also had a higher gesture rate than control children (mean difference = 0.28, 95% CI [0.07, 0.48]). This pattern suggests that explicitly asking children to gesture (*show*) increased gesture rate relative to the no instruction control.

A second explanation for the patterns we observed is that children in the three conditions differed in age, story comprehension, or performance on the verbal or spatial measures, despite assignment to condition being random. These individual differences may have been responsible for differences in CVPT gesture or narrative structure score. Because of our relatively small sample size, we elected to group particular variables for analysis rather than attempting to account for all at once (see 'Appendix'). A one factor (condition) ANOVA with age as the dependent variable showed no relationship between age and condition ($F(2,41)$ = 0.01, p = .98). A one factor (condition) ANOVA showed no relationship between condition and mean proportion correct on the comprehension questions ($F(2,41)$ = 81, p = .45). Finally, a one factor (condition) MANOVA with proportion correct on the syntax and spatial measures as the dependent variables showed no effect of condition on these measures ($F(4,82)$ = 1.66, p = .18).

The final piece of the picture we wanted to explore was whether differences in first person speech also occurred across conditions. That is, did asking children to pretend to be a character also result in changes in the extent to

423

which they used the first person to describe the events? A one factor (condition) ANOVA with first person speech as the dependent variable showed an effect of condition on first person speech ($F(2,41) = 4.47$, $p = .02$, $\eta_p^2 = .18$). Planned post-hoc comparisons showed that the CVPT children had higher rates of first person speech compared to control (mean difference = 3.66, 95% CI [0.42, 6.89]), but not OVPT (mean difference = 0.86, 95% CI [−2.3, 4.05]). OVTP and control also did not differ statistically (mean difference = 2.79, 95% CI [−0.43, 6.03]).

3.2. RECALL SESSION

As noted in the 'Method' section, recall data were lost for four children (two control, two CVPT) due to hard disk failure. Because all recall values were missing for these children, they are not included in the analysis. Our key research question is whether the training received in the narrative session impacted children's gestures and narrative structure scores in the recall session. That is, did the training condition have an effect on gesture and on narrative structure scores in the recall session? Figures 6a and 6b show the proportion CVPT gesture and the mean narrative structure scores for the recall session.

We used a one factor (condition) MANOVA with proportion CVPT and mean narrative structure score as outcome variables to explore these results. We found no effect of condition ($F(4,74) = 2.1$, $p = .09$). We next carried out a one factor ANOVA (condition) with recall gesture rate (number of gestures / number of clauses) as the dependent variable to see if differences in overall gesture rate varied. There was no significant difference in gesture rate across conditions ($F(2,37) = 1.04$, $p = .36$). Analyses for the narrative session had
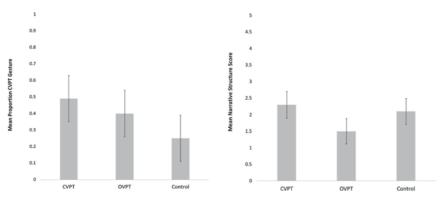

Fig. 6. Recall proportion CVPT gesture (left) and mean narrative structure score (right) by condition; error bars show standard error of the mean.

424

already established that age, story comprehension, and verbal and spatial skills did not differ across conditions. A one factor (condition) ANOVA with first person speech as the dependent variable showed no effect of condition on first person speech ($F(2,37) = 0.69$, $p = .50$). The last question to address is whether the training condition affected the number of details recalled correctly. A one factor (condition) ANOVA with overall recall as the dependent variable showed no effect of condition on recall ($F(2,37) = 1.14$, $p = .33$). In summary, these analyses indicate that the training received in the narrative session did not impact children's behavior in the recall session.

4. Discussion

Our key findings are that children who received training asking them to tell, show, and pretend to be a character produced more CVPT gestures. They also had higher narrative structure scores, though only relative to children who received the OVPT training (*tell* and *show*). Children who received training asking them to gesture (both CVPT and OVPT conditions) produced higher rates of gesture, but only those asked to 'pretend' (CVPT) had higher rates of CVPT gesture, and higher narrative structure scores. They also had higher rates of first person speech. Gesture rate, age, comprehension of stories, and performance on a language and spatial measure did not differ by condition. After a week's delay, the training did not impact gestural behavior, narrative structure scores, first person speech rates, or recall.

Why did being asked to pretend result in better narrative structure scores? One possibility is that the instruction (e.g., *Pretend to be the rabbit*) encouraged children to make the character the focus of the narration. Thus, the higher narrative structure scores might be a function of focusing the narrative on a particular actor. However, we think this is unlikely, as children in all conditions were prompted to pay attention to a specific character before describing the story. Alternatively, children who were asked to pretend may simply have been having more fun, and greater motivation or engagement might account for their higher scores. We cannot eliminate this possibility with the measures we collected. The more theoretically interesting possibility is that gesturing from the point of view of a character may have improved narrative structure because of something special about CVPT gestures. If CVPT gesture can be used to infer that the child is imagining the event from the perspective of the character, this actor perspective may result in greater focus on that character's goals and attempts. We schematized this relationship in Figure 2. In this model, greater focus on goals and attempts is the cause of higher narrative structure scores. We did see higher rates of first person speech in the CVPT condition, which offers some support for the argument that children are taking the point of view of a character. However, the better narrative structure

425

scores we observed were only in comparison to the OVPT condition, not the no instruction control. This means it would be equally valid to conclude that OVPT gesture decreases narrative structure. The claim that CVPT gesture reflects a first person conceptualization is difficult to test. Our results offer some support, but it is indirect. A next step could be to examine different types of CVPT gesture. CVPT gestures are not all equivalent (Parrill & Stec, 2018). A gesture with an actor's handshape is usually considered sufficient to code a gesture CVPT, but this handshape may reflect minimal perspective taking. A child who not only uses the same handshape as the character, but who also enacts the character's gaze and posture, mimics the character's intonation, and uses the personal pronouns and tense (e.g., first person, present tense) of the original event could be expected to have truly taken the character's point of view.

Why did we not find that CVPT gesture benefited recall a week later? Brunyé and colleagues (Brunyé et al., 2009; Brunyé et al., 2011, Ditman et al., 2010) did find that a first person encoding enhanced recall (or affected mood). Of course, they did not study gesture, but others have found gesture to benefit recall. This discrepancy may be related to an important difference between our approach and those used in previous studies. Stevanoni and Salmon (2005) found that children who were instructed to gesture at the time of recall remembered more, whereas we instructed our children to gesture at the time of encoding. On the other hand, Cook, Yip, and Goldin-Meadow (2010) instructed participants to gesture at the time of encoding, and DID find a beneficial effect on recall. Their participants were adults, so for children in this age group, it may simply be that the brief intervention was not sufficient to boost recall.

We also wonder whether intervening in children's spontaneous gesture production (by training them to produce particular kinds of gesture) might have a different effect than simply asking them to gesture, as some previous 'gesture instruction' studies have done. While being asked to gesture in particular ways can promote learning (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Cook & Goldin-Meadow, 2006; Goldin-Meadow, 2007; Goldin-Meadow & Alibali, 1995; Goldin-Meadow, Levine, Zinchenko, Yip, Hemani, & Factor, 2012; LeBarton et al., 2015; Novack et al., 2014; Ping & Goldin-Meadow, 2008; Singer & Goldin-Meadow, 2005), previous studies have examined the acquisition of principles such as mathematical equivalency and mental rotation. These concepts may not resemble the acquisition of complex narrative skills, which rely on memory, attention, linguistic ability, etc. (Berman & Slobin, 1994; Colletta, 2009; Demir et al., 2014).

What role did first person speech play in increasing narrative structure scores? The Demir et al. (2014) longitudinal study suggested that using first person speech at age five did not predict better narrative structure at later

426

ages, whereas using CVPT gesture did. We did see higher rates of first person speech in the CVPT condition, but it is important to note that children were only trained to gesture from the point of view of a character, not to adopt that character's linguistic point of view.

The role of CVPT gestures might be specific to the developmental stage we focused on. We studied kindergarteners, who are in a transitional stage in narrative development. With age, the nature of children's iconic gestures changes. Younger children produce more CVPT gestures than adults, and fewer dual viewpoint gestures (McNeill, 1992) than adults (Reig Alamillo, Colletta, & Guidetti, 2013). With age, CVPT gestures decrease, and OVPT and dual viewpoint gestures increase. Another open question is whether and how the possible facilitating role of CVPT gestures changes with age.

Whatever the underlying causes, this brief intervention did boost narrative skill. This finding might have educational implications. Prior studies aiming to improve narrative skills in children have focused primarily on children's speech. Focusing on speech reveals only a limited view of children's understanding of narratives. More extended interventions or instructions that target children's gestures as well as their speech might meet with greater success. This intervention is relatively easy to incorporate into a kindergarten curriculum. It is possible that pretending simply increases motivation rather than changing how children conceptualized the event, but an increase in motivation is hardly problematic if it leads to better narrative structure.

5. Conclusions

If children who spontaneously produce CVPT gestures at age five go on to tell better structured stories at later ages, does asking children to produce CVPT gestures improve their narrative structure immediately after? If so, the effect of embodying a character while describing its actions may change encoding in ways that have effects on narrative structure, and as has been previously stated, could subsequently facilitate literacy skills at later ages. Our study found some support for a causal role of CVPT gesture: children trained to produce CVPT gestures produced more of those gestures and had higher narrative structure scores. We take the results to indicate that gesturing from the point of view of a character promotes perspective taking. Gesture provides a crucial window into the cognitive processes that give rise to language, and these findings suggest that asking a child to take the point of view of a character can impact macro-level narrative structure. While we did not find any lasting effects on narrative structure scores, or any benefit of gestural training on recall, this was a very brief one-time intervention. Training children in CVPT story-telling repeatedly could potentially show

427

a stronger effect. This is a relatively easy classroom intervention to conduct, and thus has real practical value. Our study adds to the body of literature showing that speech and gesture inter-relate and develop in tightly coordinated ways, and showcases how gesture can be used as a tool to impact cognition and learning.

REFERENCES

Alibali, M. W., Evans, J. L., Hostetter, A. B., Ryan, K., & Mainela-Arnold, E. (2009). Gesture–speech integration in narrative: Are children less redundant than adults? *Gesture*, **9**(3), 290–311.

Applebee, A. N. (1978). *The child's concept of story: ages two to seventeen*. Chicago: University of Chicago Press.

Bamberg, M., & Marchman, V. (1990). What holds a narrative together? The linguistic encoding of episode boundaries. *Papers in Pragmatics*, **4**(1/2), 58–121.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: a crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, **136**(4), 539–550.

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives. *Psychological Science*, **20**(1), 27–32.

Brunyé, T. T., Ditman, T., Mahoney, C. R., & Taylor, H. A. (2011). Better you than I: perspectives and emotion simulation during narrative comprehension. *Journal of Cognitive Psychology*, **23**(5), 659–666.

Cassell, J., & McNeill, D. (1991). Gesture and the poetics of prose. *Poetics Today*, **12**(3), 375–404.

Colletta, J. M. (2009). Comparative analysis of children's narratives at different ages: a multimodal approach. *Gesture*, **9**(1), 61–96.

Colletta, J. M., Guidetti, M., Kunene, R., Capirci, O., Cristilli, C., Demir, Ö. E., & Levine, S. C. (2015). Effects of age and language on co-speech gesture production: an investigation of French, American and Italian children's narratives. *Journal of Child Language*, **42**(1), 122–145.

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, **84**(6), 1863–1871.

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development*, **7**(2), 211–232.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal of Memory & Language*, **63**, 465–475.

Crasborn, O., & Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. In O. Crasborn, T. Hanke, E. Efthimiou, I. Zwitserlood, & E. Thoutenhoofd (Eds.), *Construction and exploitation of sign language corpora: 3rd workshop on the representation and processing of sign languages* (pp. 39–43). Paris: ELDA.

Dancygier, B. (2012). *The language of stories: a cognitive approach*. Cambridge: Cambridge University Press.

Debreslioska, S., Özyürek, A., Gullberg, M., & Perniss, P. (2013). Gestural vewpoint signals referent accessibility. *Discourse Processes*, **50**, 431–456.

Demir, Ö. E., Levine, S. C., & Goldin-Meadow, S. (2014). A tale of two hands: children's early gesture use in narrative production predicts later narrative structure in speech. *Journal of Child Language*, **42**(3), 662–681.

Ditman, T., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: pronouns guide action simulation during narrative comprehension. *Cognition*, **115**, 172–178.

Feagans, L. V., & Short, E. J. (1984). Developmental differences in the comprehension and production of narratives by reading disabled and normally achieving children. *Child Development*, **55**, 1727–1736.

Gillam, R. B., & Johnston, J. R. (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech and Hearing Research*, **35**(6), 1303–1315.

Goldin-Meadow, S. (2005). The two faces of gesture: language and thought. *Gesture*, **5**(1/2), 241–257.

Goldin-Meadow, S. (2007). Pointing sets the stage for learning language — and creating language. *Child Development*, **78**, 741–745.

Goldin-Meadow, S., & Alibali, M. W. (1995). Mechanisms of transition: learning with a helping hand. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 115–157). San Diego, CA: Academic Press.

Goldin-Meadow, S., Levine, S. C., Zinchenko, E., Yip, T. K. Y., Hemani, N., & Factor, L. (2012). Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science*, **15**(6), 876–884.

Griffin, T. M., Hemphill, L., Camp, L., & Palmer Wolf, D. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, **24**(2), 123–147.

Halliday, M., & Hasan, R. (1979). *Cohesion in English*. London: Longman.

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: relations between verbal and spatial skills and gesture production. *Gesture*, **7**(1), 73–95.

Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin*, **105**(2), 302–308.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. C. (2002). Language input and child syntax. *Cognitive Psychology*, **45**, 337–374.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, **16**(5), 367–371.

Jones, C. H., & Pipe, M.-E. (2002). How quickly do children forget events? A systematic study of children's event reports as a function of delay. *Applied Cognitive Psychology*, **16**, 755–768.

Karmiloff-Smith, A. (1979). *A functional approach to child language*. Cambridge: Cambridge University Press.

Labov, W., & Waletzky, J. (1997). Narrative analysis: oral versions of personal experience. *Journal of Narrative & Life History*, **7**(1/4), 3–38.

Langdon, R., Michie, P. T., Ward, P. B., McConaghy, N., Catts, S. V., & Coltheart, M. (1997). Defective self and/or other mentalising in schizophrenia: a cognitive neuropsychological approach. *Cognitive Neuropsychiatry*, **2**(3), 167–193.

LeBarton, E. S., Raudenbush, S., & Goldin-Meadow, S. (2015). Experimentally-induced increases in early gesture lead to increases in spoken vocabulary. *Journal of Cognition and Development*, **16**(2), 199–220.

McCabe, A., & Peterson, C. (1991). *Developing narrative structure*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Nicoladis, E., Marentette, P., & Navarro, S. (2016). Gesture frequency linked primarily to story length in 4–10-year old children's stories. *Journal of Psycholinguistic Research*, **45**(2), 189–204.

Ninio, A., & Snow, C. E. (1996). *Pragmatic development*. London: Routledge.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: using the hands to learn math. *Psychological Science*, **25**(4), 903–910.

Özçalışkan, Ş., & Goldin-Meadow, S. (2009). When gesture–speech combinations do and do not index linguistic change. *Language and Cognitive Processes*, **24**(2), 190–217.

Parrill, F. (2010). Viewpoint in speech–gesture integration: linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes*, **25**(5), 650–668.

429

Parrill, F., & Stec, K. (2018). Seeing first person changes gesture but saying first person does not. *Gesture*, **17**(1), 159–176.

Petersen, D. B. (2011). A systematic review of narrative-based language intervention with children who have language impairment. *Communication Disorders Quarterly*, **32**(4), 207–220.

Peterson, C., Jesso, B., & McCabe, A. (1999). Encouraging narratives in preschoolers: an intervention study. *Journal of Child Language*, **26**(1), 49–67.

Ping, R., & Goldin-Meadow, S. (2008). Hands in the air: using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology*, **44**(5), 1277–1287.

Reig Alamillo, A., Colletta, J.-M., & Guidetti, M. (2013). Gesture and language in narratives and explanations: the effects of age and communicative activity on late multimodal discourse development. *Journal of Child Language*, **40**, 511–538.

Reilly, J. S. (1992). How to tell a good story: the intersection of language and affect in children's narrative. *Journal of Narrative and Life History*, **2**(4), 355–377.

Singer, M., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, **16**(2), 85–89.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 53–120). Norwood, NJ: Ablex.

Stevanoni, E., & Salmon, K. (2005). Giving memory a hand: instructing children to gesture enhances their recall. *Journal of Nonverbal Behavior*, **29**(4), 217–233.

Thurstone, L. L. (1974). *The measurement of values*. Chicago, IL: University of Chicago Press.

Warden, D. (1976). The influence of context on children's use of identifying expressions and references. *British Journal of Psychology*, **67**, 101–112.

Wesson, M., & Salmon, K. (2001). Drawing and showing: helping children to report emotionally laden events. *Applied Cognitive Psychology*, **15**, 301–320.

Appendix

Materials: cartoons used

Baseball Bugs (Bugs Bunny, Freleng, 1946), *Fast and Furrious* (Wile E Coyote, Jones, 1949), *For Scent-imental Reasons* (Pepe Le Pew, Jones, 1949), and *Feed the Kitty* (Marc Antony, Jones, 1952)

Sample training gestures

See Table A1.

Syntax comprehension task

In the syntax comprehension task (Huttenlocher et al., 2002), the experimenter reads 54 sentences to the child, each describing a single or multi-clause event. Sentences increase in complexity over the course of the task (e.g., sentence 1: *The boy is eating pizza*; sentence 54: *A girl who is standing next to the boy is holding at kite*). Children see a set of pictures and have to choose the picture that matches the sentence. Figure A1 shows the pictures paired with sentence 54 (the example given above). For questions 11–55, one picture was intended as a 'foil', that is, it could be interpreted as correct if the child did not attend

TABLE A1. *Sample training gestures*

Speech produced in all three conditions: A mouse is standing on a shelf in a kitchen and holding onto a rope and then he swings down from the shelf!	
Character viewpoint	1. Character viewpoint two hands hold rope.
	2. Character viewpoint two hands hold rope + move away from body on sagittal axis.
Observer viewpoint	1. Observer viewpoint one hand traces rope up to down
	2. Observer viewpoint one hand traces path of mouse on transverse axis left to right

NOTE: Gestures produced in one of the training trials, with descriptions of the CVPT and OVPT versions. Gestures occurred during the underlined speech. In the control condition, the underlined speech was produced with no gesture.
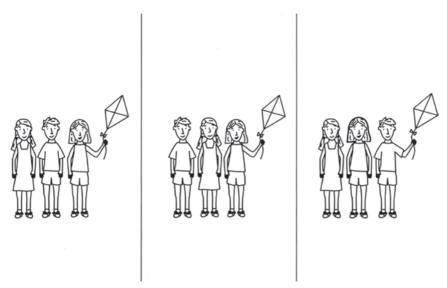

Fig. A1. Syntax comprehension sample item.

carefully to the syntax (the middle picture in Figure A1). This task is a basic measure of oral language.

Spatial measure task

In the Thurstone Primary Mental Abilities Spatial Relations Subtest (Thurstone, 1974), children see a square with a shape missing and must choose which of four possibilities would complete the square. This subtest was done with paper and pencil. Figure A2 shows a sample item.

431

Fig. A2. Spatial test sample item.

Training: prompts for gesture conditions

When children did not produce the correct gesture, the experimenter followed this set of prompts:

1. CHARACTER VIEWPOINT CONDITION. *I like how you told me, but you forgot to pretend you were the* [character]*! Remember when I said* [repeat second speech–gesture pair from training]. *I was pretending to be the* [character]*! Now can you tell me and show me what happened and pretend you're the* [character]*?*
2. OBSERVER VIEWPOINT CONDITION. *I like how you told me, but you forgot to show me! Remember when I said* [repeat second speech–gesture pair from training]*? I was showing you what happened! Now can you tell me and show me what happened?*

Comprehension questions

What did the rabbit hit?
Why did the rabbit ride a bus?
Did the rabbit ride on an elevator?
Who was the skunk following?
Why was skunk following that animal?
Did the skunk run upstairs?
What kind of an animal did the dog bring home?
Why did the dog try to hide the animal?
Did the dog climb on top of a ball?
What did the coyote have around his neck?
Why did the coyote have a knife and fork?
Did the coyote chase a mouse?

Narrative structure scoring examples

Descriptive sequence (1): *The skunk wanted to give the cat love*. Scored as descriptive due to lack of temporal organization.
Action sequence (2): *The dog set up a cat thing and it ran into the person. The kitty was on a boot. That's all I have to tell*. Scored as action sequence because had temporal but no causal structure.

432

Reactive sequence (3): *What happened was um his dad chased him. And then he went inside the room and then he pushed the window out to see and then he said hello.* Scored as reactive because had causal structure, but no goal.

Incomplete goal-based sequence (4): *What happened was he climbed up and then he throwed his glove to catch the ball.* Scored as incomplete because had goal but no outcome.

Complete goal-based story with one episode (5): *There was a coyote chasing the ostrich and the ostrich got away because it was too fast and the wolf the um coyote was too tired.* Scored as complete: had both goal and outcome.

Demographic data

See Table A2.

Descriptive statistics

See Table A3.

Statistical analyses

Analyses were carried out using IMB SPSS Statistics version 24. The number of participants in this study is small, while the number of outcome variables is large, limiting the kinds of analyses we could perform. Our goal was to perform the simplest analyses that would answer our main research questions and explore obvious alternative explanations, were we to find an effect of training on our key outcome variables. We sometimes conducted ANOVA rather than MANOVA in order to preserve the ability to detect differences. Grouping of variables was guided by conceptual considerations rather than by inter-correlations (Huberty & Morris, 1989).

TABLE A2. *Demographic data*

Category	Percentage
Native English speakers	97
Speak a language other than English at home at least 10% of the time	23
Mother with post-graduate degree	50
Mother with BA/BS	29
Father with post-graduate degree	53
Father with BA/BS	15
Race/Ethnicity	European American: 74, Asian: 16, Unknown: 8, African American: 2

TABLE A3. *Descriptive statistics for variables of interest (SDs in parentheses)*

	CVPT	OVPT	Control
Age	6:1 (0.38)	6:1 (0.42)	6:1 (0.51)
Gesture rate	0.36 (0.19)	0.40 (0.27)	0.12 (0.16)
Proportion CVPT	0.88 (0.12)	0.46 (0.25)	0.39 (0.42)
Mean narrative structure score	3.42 (1.04)	2.54 (0.87)	2.91 (0.77)
First person speech	6.73 (4.17)	5.87 (3.71)	3.07 (1.86)
Proportion correct comprehension	0.81 (0.11)	0.79 (0.13)	0.85 (0.12)
Proportion correct syntax	0.85 (0.06)	0.83 (0.07)	0.89 (0.07)
Proportion correct PMA spatial	0.70 (0.13)	0.70 (0.23)	0.79 (0.16)
Recall gesture rate	0.12 (0.19)	0.18 (0.13)	0.19 (0.18)
Recall proportion CVPT	0.49 (0.38)	0.40 (0.36)	0.25 (0.33)
Recall mean narrative structure score	2.3 (1.16)	1.5 (1.13)	2.09 (0.42)
Recall first person speech	10.46 (7.37)	8.2 (3.29)	9.58 (3.91)
Recall overall	17.08 (6.27)	13.6 (8.19)	17 (5.88)

Assumption testing for (M)ANOVA

Sample size rules of thumb vary – we used the guideline that the number of observations should be larger than number of levels multiplied by number of dependent variables. We also choose not to analyze more than two dependent variables at the same time to avoid violating sample size rules of thumb. Extreme values were assessed using stem and leaf plots or Mahalanobis distances. Maximum Mahalanobis distances did not exceed the critical value in any case. There were indeed extreme values, and we verified the scoring on all to ensure there were no errors in the data. These extreme values represent real variability on some of these measures (e.g., particularly high or low spatial scores). Multivariate normality was assessed using the Shapiro–Wilk statistic. This assumption was not always met. In cases where all assumptions were met, the statistics reported use Wilks' Lambda, and when not met, Pillai's trace. Linear relationships among dependent variables were assessed using scatterplots. Homogeneity of variance was tested using Levene's test of equality of error and non-significant results occurred in all cases. Homogeneity of covariance was tested using Box's test of equality of covariance matrices. Box's M values exceeded the cut-off value of .001 in all cases. Multicollinearity was assessed using bivariate correlations. In all cases, we found a relationship that exceeded the minimum (using a .2 cut-off) and did not exceed the maximum (using a .8 cut-off). Planned post-hoc comparisons used Scheffés test because of unequal group sizes and robustness to non-normality.