IN PRESS: DEVELOPMENTAL PSYCHOLOGY

Handling agents and patients: Representational co-speech gestures help children comprehend complex syntactic constructions.

Anna L. Theakston[1]

Anna Coates[1]

Judith Holler[1,2]

[1] *University of Manchester, UK*

[2] *Max Planck Institute for Psycholinguistics, NL*

Address for correspondence: Dr. Anna Theakston, School of Psychological Sciences, University of Manchester, UK. Email: anna.theakston@manchester.ac.uk

Acknowledgements

Thanks are due to the schools, nurseries, children and families who took part in the study, as well as to three anonymous reviewers who provided invaluable feedback on an earlier version of this manuscript. JH was supported by ERC Advanced Grant #269484-INTERACT during parts of this project. Parts of the data collection were funded by the Max Planck Institute for Evolutionary Anthropology.

Abstract

Gesture is an important precursor of children's early language development, for example, in the transition to multiword speech and as a predictor of later language abilities. However, it is unclear whether gestural input can influence children's comprehension of complex grammatical constructions. In Study 1, three- (M = 3;5) and four-year-old (M = 4;6) children witnessed two-participant actions described using the infrequent object-cleft-construction (OCC- *It was the dog that the cat chased*). Half saw an experimenter accompanying her descriptions with gestures representing the two participants and indicating the direction of action, the remaining children did not witness gesture. Children who witnessed gestures showed better comprehension of the OCC than those who did not witness gestures, both in and beyond the immediate physical context, but this benefit was restricted to the oldest four-year-olds. In Study 2, a further group of older four-year-old children (M = 4;7) witnessed the same two-participant actions described by an experimenter and accompanied by gestures, but the gesture represented only the two participants and not the direction of the action. Again, a benefit of gesture was observed on subsequent comprehension of the OCC. We interpret these findings as demonstrating that representational co-speech gestures can help children comprehend complex linguistic structures by highlighting the roles played by the participants in the event.

Key words: acquisition of grammatical constructions, object-cleft-construction, co-speech gestures, learning, syntax

There is a growing interest in the role that gesture plays in children's early communicative development. Researchers have investigated child and maternal gesture use as a predictor of children's later language abilities, demonstrating that early advantages in the gestural modality precede early advances in spoken language at the very earliest stages of language acquisition. However, very little work has examined whether the gestures children are exposed to might act as a scaffold for the comprehension of complex linguistic constructions. Here, we report two studies examining the role of representational co-speech gestures in children's comprehension of a complex (object cleft) syntactic construction.

Children's gesture production and subsequent language development

A body of evidence demonstrates the predictive value of children's own gestures for their subsequent language acquisition. In typical development advances in children's use of gesture typically precede developments in their spoken language predicting their vocabulary size (Acredolo & Goodwyn, 1988; Brooks & Meltzoff, 2008; Carpenter, Nagell, & Tomasello, 1998; Colonesi, Stams, Koster & Noom, 2010) and content (Iverson & Goldin-Meadow, 2005). Similar predictive relations can be found in atypical populations. Children with brain lesions whose early gesture use falls within the normal range show typical vocabulary development, whereas those who gesture below the typical range show delays in subsequent vocabulary development (Sauer, Levine & Goldin-Meadow, 2010).

Gesture is also a precursor to and predictor of the transition to multiword speech. Children start to produce supplementary gesture-speech combinations (the word and gesture convey different but related concepts) at around 17-months, approximately four months before the emergence of spoken two-word combinations (Capirci, Contaldo, Caselli, & Volterra, 2005; Iverson & Goldin-Meadow, 2005, see Özçalışkan & Goldin-Meadow, 2010 for work on gender differences in gesture and the onset of multiword speech). This predictive relation holds independently for different grammatical constructions (Özçalışkan

& Goldin-Meadow, 2009), and in addition, the number of gesture+speech combinations produced at 18-months correlates with sentence complexity at 42-months (Rowe & Goldin-Meadow, 2009). Thus, an early start in the gestural modality is associated with advantages in the spoken modality up to at least three-years of age.

Children’s comprehension of gestural input

Although children’s own gestural production has associations with their later language learning, the key focus in this paper is on the relationship between gestural input and children’s language comprehension. Prior to three-years, children are exposed to what has been termed ‘gestural motherese’ (Iverson, Capirci, Longobardi & Caselli, 1999). The majority of spontaneous gestures produced by caregivers interacting with 16-20-month-old children (around 70%) are deictics and conventional gestures that are firmly situated within the immediate context, with very little use of representational or emphatic gestures, or gestures referring to more abstract concepts or absent referents (Bekken, 1989; Iverson, et al., 1999; Rodrigo, Gonzales, Ato, Rodriguez, de Vega, & Muneton, 2006). The exact profile of maternal gesture use may change according to task demands (Grimminger, Rohlfing & Stenneken, 2010), and there is an increase in maternal use of iconic gestures when children are around 26-months old (Özçalışkan & Goldin-Meadow, 2011).

An important question related to children’s gestural input is whether they are able to glean information from their caregivers’ gestures, and to what extent this improves their language comprehension. Adults can extract information from the gestures of their interlocutors, especially those that relate to spatial and/or motor information (e.g., Beattie & Shovelton, 1999; Graham & Argyle, 1975; Holler, Shovelton & Beattie, 2009; Kelly, Barr, Church & Lynch, 1999; Riseborough, 1981; see Hostetter, 2011 for a review). For children, there is considerably less research on this issue. We do know that between one and two years of age children begin to interpret pointing gestures as referentially intended (Behne, Carpenter & Tomasello, 2005), map symbolic gestures to their referents (Namy, Campbell & Tomasello, 2004; Tomasello, Striano, & Rochat, 1999, also Marentette & Nicoladis, 2011), and interpret bimodal messages encoded by a simple supplementary gesture (hand outstretched=GIVE) presented alongside a spoken single word ('Ball', Morford & Goldin-Meadow, 1992). Thus, where gestural input has been shown to be helpful to children in understanding communicative intent and in facilitating language comprehension and production, the advantages seen are largely based on deictic or iconic gestures and concern relatively simple propositions.

Relatively little research examines whether older children can use gestures to assist in their comprehension of more complex verbal messages, and whether the nature of the gestures children are able to interpret becomes more abstract (see Kelly, 2001, for the role of pointing gestures in 3-5-year-olds' comprehension of indirect requests). McNeil, Alibali and Evans (2000) tested 3-6-year-olds' comprehension of complex spoken messages (e.g. *Find the block that has an arrow pointing up and a smile face with a rectangle above it*) accompanied by reinforcing gestures depicting the direction of the arrow, the relative positioning of the two components on the target block and, the shape of the referents. They found that the 3-4-year-olds (M=51-months) benefitted from these gestures, whereas the 5-6-year-olds (M=64-months) could rely on their linguistic skills alone. On the other hand, only the older children were confused when the gestural input contradicted information in the spoken message, suggesting that children's ability to make use of different kinds of gestural input, and their ability to integrate this information with speech, changes over development. In this study, however, the gestures depicted concrete aspects of the visual scene, and the language task required the comprehension of spatial terms (e.g. *up, down, above, below*). This differs from the interpretation of semantic roles such as agent and patient, as is required in the comprehension of most grammatical constructions. It is therefore unclear whether children would benefit from more abstract co-speech gestures, whether gestural input can assist children in the comprehension of grammatical structures, and how this might change over the course of development. The studies we report here are a first attempt in advancing our understanding in this domain.

Mechanisms for syntax acquisition

By three-years of age, children have begun to produce the majority of syntactic structures for their language, but their comprehension and production of complex structures such as relative clauses and questions with long distance dependencies is not fully adult-like (Correa, 1995; Dabrowska, Rowland & Theakston, 2009). In particular, although even two-year-olds demonstrate an understanding of participant roles in simple active transitive sentences (Gertner, Fisher & Eisengert, 2006), children have more difficulties disentangling participant roles in noncanonical sentence frames in which the patient of an action appears before the agent, for instance in object relative clauses with two lexical nouns (Brandt, Kidd, Lieven & Tomasello, 2009) and English passives (Bever, 1970; Gordon & Chafetz, 1990).

In the current study we are interested in the role that co-speech gestures, more specifically, representational gestures (Alibali et al., 2001), might play in conveying information about different participants in an event during children's comprehension of complex sentences about such events. Representational gestures can identify referents by depicting particular features of the respective referent iconically through the shape of the hand(s), and/or by means of local association; in narrative, for example, speakers tend to associate individual referents with particular locations in their gesture space to track the referents and the relations amongst them (Gullberg 2006; So, Kita & Goldin-Meadow, 2009). In some cases, speakers combine these forms of representation by localising iconically specific gestures, while, in others, they perform 'neutral' hand shapes that do not depict any particular characteristics of the referents. Instead, the hand adopts a neutral form (e.g., slightly cupped hand shape, palm pointing downwards), to indicate the existence of an entity in a particular location in space (speakers also use pointing gestures to locate entities in gesture space, 'abstract pointing', McNeill, 1992). The present research focuses on representational gestures of this 'form-neutral' kind in the context of events involving one patient and one agent (henceforth 'entity gestures'). By using this more abstract (or metaphorical) kind of representational gestures, we build on earlier studies of iconic gesture comprehension (Namy et al. 2004, Namy, 2008; Namy & Waxman, 1998; Stanfield, Williamson & Ozçalişkan, 2014)

The rationale for assuming that these kinds of co-speech gestures might help children in their comprehension of non-canonical syntactic structures is threefold. First, children's learning of gestures and spoken language appears to go hand-in-hand, suggesting that the same underlying mechanisms are involved, or that gesture relies on 'knowledge relevant to the process of learning language' (Ozcaliskan & Goldin-Meadow, 2010:753). This fits with the theory that spoken language and co-speech gestures form part of an integrated multimodal communication system (e.g., Kendon, 2004; McNeill, 1992), are learned together, and are used together to work out a speaker's meaning. Since there is evidence that deictic and iconic gestures benefit children's comprehension of spoken language (as seen above), we might expect that gestures of a more abstract, metaphoric nature may improve at least older children's comprehension of spoken messages, too. Second, in other areas of cognitive development, co-speech gestures have been shown to benefit children. 5-10-year-old children are able to detect information presented in co-speech gestures in others' explanations of Piagetian conservation tasks (Kelly & Church, 1997, 1998), and perform better on these tasks (Ping & Goldin-Meadow, 2008), and on mathematical equivalence problems (e.g. Goldin-Meadow, Kim, & Singer, 1999; Singer & Goldin-Meadow, 2005) if they are given instruction containing gestures. This is argued to reflect a reduction in performance demands such that, on tasks that depend on knowledge that children are only just acquiring, gesture serves to reduce the cognitive load. Similarly, children may also benefit from a reduced cognitive load in language comprehension at the age at which they have only a partial knowledge of a particular grammatical construction. Finally, understanding syntactic structures depends on understanding what role each participant plays in an event. Therefore, gestures that mark the semantic roles of the referents, such as the abstract entity gestures outlined above, could be particularly helpful in this context.

Linguistically, the focus of the current study is the object-cleft-construction (OCC), for example *It was the [object] frog that the [subject] man pushed*. This construction is expected to be difficult for children to comprehend and produce because it contains exactly the same words as the related subject-cleft-construction (*It was the [subject] frog that pushed the [object] man*), thus the only cue to participant roles is the syntactic construction[1]; it is sometimes misinterpreted by adults because the ordering of participant roles is noncanonical, i.e. the patient (object) appears before the agent (subject) (Ferreira, 2003); and it is late acquired. Even 6-year-olds are below ceiling (69%) in their correct interpretation of object-cleft sentences in an act-out task (Lempert & Kinsbourne, 1978), and French-speaking four-year-olds rarely spontaneously produce full object-clefts (<3%), even in pragmatically appropriate contexts (Hupet & Tilmant, 1989). Ambridge, Theakston, Lieven and Tomasello (2006) attempted to teach 3-5-year-olds object-cleft-constructions with an animate agent and inanimate patient, and found that in their massed condition (training occurred on a single day analogous to the current study), when given four opportunities to produce a novel object-cleft, only 27% of the 3-4-year-olds' ($\underline{M}_{age}$ 4;3) and 35% of the 4-5-year-olds' ($\underline{M}_{age}$ 5;4) responses were object-cleft-constructions, with only five of twelve children in each group

[1] compare with a passive sentence such as *The frog was push<u>ed by</u> the man*, where the *by*-phrase and verb morphology give important clues to the identity of the agent and patient

producing a novel object-cleft[2]. This demonstrates that this is a difficult construction for 3-5-year-olds to learn and therefore provides a good opportunity to investigate the possibility that gesture might improve the learning of complex grammatical constructions.

In Study 1, entity gestures (see above) are combined with a subsequent representational gesture indicating the *direction* of the action being performed by one of the characters—again in a form-neutral manner, namely by representing the direction of the action through a generic sweep from entity 2 (agent) to entity 1 (patient), *without depicting any characteristics of the particular action performed* (see Fig. 1). We hypothesised that even representational gestures of this rather abstract and 'bleached' nature, used to reinforce and clarify the information encoded in speech, would provide a scaffold to children's interpretations of the object-cleft-construction in utterances where both participants are animate, creating ambiguity regarding their semantic roles. Specifically, by using one hand to represent the patient, and the other to represent the agent of the action, as well as the direction of movement associated with this action (i.e., generic movement from agent to patient), gesture might help children to focus on the ordering of the participants and their roles in the non-canonical object-cleft-construction. However, abstract representational gestures indicating the existence of referents and the relations among them are, despite their simplicity in shape, relatively complex constructs; children need to understand that these gestures stand in for the characters represented and the actions they carry out, even though their form bears no physical resemblance to the characters or the particular actions performed. Although it could be argued that these rather complex abstract gestures may not be typical of adults' communication directed at young children, it is important to bear in mind that in much of the daily communication children experience they are in an overhearer role (and it has been

[2] In their distributed-pairs condition, learning was significantly improved with eight of twelve 3-4-yr-olds and twelve of twelve 4-5-yr-olds producing at least one object-cleft, with object-clefts accounting for 39% and 94% of their responses respectively, i.e. 4-5-yr-olds showed the most improvement according to training condition.

shown experimentally that children learn language and even comprehend gestures directed at third persons, e.g. Akhtar, Jipson, & Callanan, 2001; Gräfenhain, Behne, Carpenter, & Tomasello, 2009). As such, they may frequently see, and thus learn through, entity gestures (as well as other more abstract gestures) that may be primarily directed at older children or other adults.

While there is evidence that children's understanding of the use of highly iconic co-speech gestures undergoes critical changes between the ages of three and four (such as in the context of homonymy, Kidd & Holler, 2009), it is unknown whether similar developmental changes exist for gestures of a more abstract - and in this case also fairly bleached - nature. The studies reported here address this question.

Study 1

Method

Participants. 86 monolingual English-speaking children participated in the study. There were 41 three-year-olds (19 girls, 3;1-3;11, M_{age} = 3;5 in both the gesture and non-gesture groups) and 45 four-year-olds (28 girls, 4;3-4;8, M_{age} = 4;6 in both the gesture and non-gesture groups). A further 7 children (6 girls) were excluded because they scored at ceiling on both a baseline comprehension task and during the training phase (*n* = 4, by chance only 1/256 children would show this pattern of behaviour), and thus did not have the capacity to show improvement in performance as a result of manipulations in input administered during the study, and due to experimenter error in recording of date of birth or condition information (*n* = 3). The children were recruited from nurseries and preschools in the Manchester and Sheffield areas of the UK, and through the Child Study Centre at the University of Manchester.

Materials and Stimuli. The grammatical construction selected for use was the object-cleft-construction (OCC, *It was the [object] dog that the [subject] cat pushed*). Twenty-four transitive causative verbs appropriate for use to describe the action of one puppet on another were chosen from the Wells corpus (Wells, 1981) available on the CHILDES website (MacWhinney, 2000), which contains transcripts of speech from children aged 1;6-5;0 (see Appendix A).

Twenty-four silent video clips, each showing two hand puppets engaged in one of the 24 transitive actions, were created. Ten puppets were used in the videos, and all were easily identified by the participants. Each puppet appeared in four or five clips in total, at least twice as the agent and twice as the patient of the action. Puppets were never paired with each other more than once. The positioning of the agent of the action on the left and right of the screen was counterbalanced.

For each action clip, a corresponding video clip was created of the Experimenter (E) describing the action using an object-cleft-construction. This clip was shown twice in succession, thus each action was described twice. Hand gestures were used alongside the spoken description. To determine the kinds of gestures normally produced when describing these scenes, 10 adults were given examples of the object-cleft-construction and asked to watch the videos and describe the actions using an object-cleft-construction to another adult who had not witnessed the videos. The adults predominantly produced entity gestures of a form-neutral kind (such as a loose but slightly cupped hand shape with the palm pointing downwards, see Introduction) when referring to the puppets and simple, uni-directional sweeping gestures to indicate the direction of motion between the characters. These gestures served as templates for the actor performing the gestures in our video clips, as in (1);

(1) *it was the [dog] that the [duck] [scratched]*

[Left-hand extends outwards at chest level, slightly cupped, with palm facing the floor and remains stationery (see Figure 1a)], then [right-hand extends outwards with palm facing the floor remains stationery (see Figure 1b)] and then [right hand sweeps horizontally towards the stationery left hand (see Figure 1c)]

FIGURE 1 HERE

An entity gesture was produced to coincide with the naming of each puppet, and a sweeping gesture to coincide with the verb denoting the action. The left hand was used to denote the patient, and the right hand to denote the agent of the action. After localising the two participants in the event, the right hand produced the sweeping movement towards the left hand (still in its original position), irrespective of the relative locations of the two puppets on the video or the actual manner of action. E was filmed against a plain white backdrop, with her upper body (including the face) captured on screen, to ensure that all hand gestures were visible to the viewer. These clips were used in the gesture condition. For the no-gesture condition, the clips were cropped to show only E’s face, and not her hand gestures. This method ensured that the video stimuli used in each condition were identical in all respects (e.g., intonation, speech rate, facial expression) except for the presence/absence of hand gestures.

To test the children’s comprehension of the object-cleft sentences, 48 picture cards were created. For each of the 24 object-cleft sentences, one picture card showed the agent acting on the patient (target - correct), the other showed the patient acting on the agent (foil - incorrect), see Figure 2. These cards also served to motivate the pragmatically appropriate use of the OCC. Typically this construction is used to focus attention on the patient of a causative action and/or to indicate a contrast between that participant and another. The two contrasting scenes ensured that children had to pay attention to the particular roles played by each character in the action, and thus licensed the use of the OCC.

FIGURE 2 HERE

Design. Participants were allocated to either the gesture or no-gesture condition. The study comprised of four comprehension tasks (baseline, training, comprehension 1, comprehension 2). Four object-cleft sentence trials were used in each task. The tasks were administered in the same order across children, but four different scripts were created such that the allocation of actions (verbs) to tasks differed across children.

In the comprehension tasks, children had to choose one of two pictures (target and foil) that corresponded to an object-cleft-construction (OCC). In the *baseline* task, children heard E producing an OCC with carefully controlled intonation (to match the subsequent videos) without any gesture (irrespective of condition). This established a baseline level of comprehension of the OCC against which subsequent performance could be compared. Note that in the baseline task, children interacted with E in person so were fully aware that she was not gesturing during her production of the OCC. This is important because it sets up an expectation, for later video trials, that E is not gesturing during her descriptions of the scenes. Thus, when children witnessed cropped videos showing only E’s face, this is unlikely to have caused any particular confusion for them. In the *training* task, children first watched a video of the action performed by two characters, then saw a video of E describing the action with/without gesture according to condition. This gave children an opportunity to learn something about the appropriate use of the construction by pairing a visual scene with the OCC, and allowed us to determine whether training was more effective with the inclusion of gesture. In the *comprehension 1* task, children saw a video of E producing an OCC with/without gesture according to condition. This allowed us to determine whether comprehension of the construction alone (without video of the action) differed as a function of the inclusion of gesture. Finally, in the *comprehension 2* task, all children irrespective of condition saw a video of E producing an OCC without any gesture. This allowed us to investigate whether children who had been exposed to OCCs with gesture could extend their knowledge of the construction to instances when gesture was no longer present.

Procedure. Children were tested individually in a quiet area of their school or nursery, or at the Child Study Centre. They sat next to the experimenter in front of a laptop computer. Children were videoed throughout to record their audio responses and to allow coding of their picture selections (another aim was to see whether the children would produce gestures themselves in the two production tasks; however, too few instances occurred in our data set to warrant any systematic analyses). In the baseline task, E explained that the children would see two picture cards with two puppets on them, and she would describe how the puppets were playing. Only one of the picture cards would match what she had said, and the children's task was to decide which one was right and point to it. Two picture cards were then placed in front of and at an equal distance from the child, and E described an action using an OCC. The positioning of the target card on the left or right was pseudo-randomised such that it never appeared on the same side more than three times in succession. Once children had pointed to a card, their response was recorded and E gave them positive encouragement. This procedure was repeated for all four trials. Children were then introduced to the video tasks. In the training task, they were told that rather than just hearing E describing what had happened between two puppets, they would also get to see for themselves and they needed to watch carefully. They were then shown an action clip, immediately followed by a clip of E describing the action using an OCC which was repeated twice (on rare occasions when the child did not attend to the video, the clip was played again). The picture cards were then placed in front of the child (to prevent children from looking at the cards during the videos and thus not seeing either the action or E's gestures if applicable). Children were encouraged to point to the card they thought corresponded to the action and were always given positive feedback (e.g. *that's great*, *well done*) to maintain their interest. This was repeated for all four trials in the training task. The same procedure was followed for the two subsequent comprehension tasks except that children were told they would not see a video of what the puppets had done so would have to listen carefully to what E said on the video to decide what had happened.

Coding. For the four comprehension tasks, for each of the four trials children received a score of one if they picked the correct card and zero if they picked the foil card. To check for inter-coder reliability, 25% of the data were coded by a second researcher resulting in 100% agreement on the selection of the target vs. foil picture cards.

Results

To determine whether hearing object-cleft sentences produced alongside representational gestures resulted in a better comprehension of the OCC during the training phase, the comprehension 1 or the comprehension 2 tasks, relative to performance at baseline, we examined the number of correct responses (selection of the correct picture card, out of 4) on each of the four tasks. Figure 3 shows the mean number of correct responses in each task, splitting the data by age for illustrative purposes (3yrs, 4yrs).

FIGURE 3 HERE

To analyse the data, we fitted a mixed effects logistic regression model using Laplace approximation, with response (correct vs. incorrect) as the outcome variable and participants and verbs as random factors (Baayen, Davidson & Bates, 2008). We first compared models with and without the random effect of verb. The model including the random predictor ‘verb’ provided a significantly better fit to the data (χ^2 (1) = 3.75, p = .05), thus both random predictors of participant and verb were retained to control for differences in response accounted for by individuals or items. Gesture condition (no gesture, gesture), age (a continuous variable measured in months, centred on the 50th percentile at age 3;9) and task (baseline, training, comprehension 1, & comprehension 2) were then entered into the model as predictor variables, along with the interactions between them. The model compared the children’s responses in each of the three tasks of interest (training, comprehension 1, and comprehension 2) with their performance at baseline to establish if and when gesture would benefit the children’s comprehension. Children’s responses in the no gesture condition on the baseline task provided the point of reference. Table 1 shows the coefficients of the model for the main experimental manipulation of gesture, for the main effects of task, and for those additional predictors that had a significant effect.

TABLE 1 ABOUT HERE

Taking the simple effects first, the data show that the children in the gesture and no-gesture conditions did not differ significantly in their performance on the baseline task. Nor did children in the no-gesture group show any significant differences in performance between the baseline task and the other tasks. This confirms that any effects of gesture reflect differences in performance once the children had been exposed to the training stimuli. There was a significant interaction (p = .02) between age and gesture, when comparing the baseline and the training tasks, suggesting that gesture has a facilitative effect on responses in the training task, but not at all ages. To investigate this interaction, additional models were fitted to the data with age centred on the 25th (age 3;5), and 75th (age 4;6) percentiles. The results show that there was a significant interaction between task (training vs. baseline) and gesture for the oldest children (75th percentile, from age 4;6: $\beta = 1.05$, $SE = 0.43$, $z = 2.47$, $p = .01$) and a marginally significant interaction for the mid-age children (50th percentile $\beta = 0.62$, $SE = 0.33$, $z = 1.88$, $p = .06$) with only those who received gestural input having improved performance on the training task compared to at baseline (interaction between task & gesture at 25th percentile $\beta = -0.23$, $SE = 0.41$, $z = -0.57$, $p = .57$).

There was also a significant interaction between age and gesture when comparing the comprehension 2 task against baseline ($p = .01$), suggesting that gesture had a facilitative effect on comprehension on this task, but not at all ages. Additional models were fitted to the data with age centred on the 25th (age 3;5), and 75th (age 4;6) percentiles to investigate the interaction. The results show that there was a significant interaction between task (baseline vs. comprehension 2) and gesture for only the oldest children (75th percentile: $\beta = 0.91$, $SE = 0.42$, $z = 2.47$, $p = .03$) with only those who received gestural input having improved performance on the comprehension 2 task compared to baseline (interaction between task & gesture at 25th percentile $\beta = -0.56$, $SE = 0.41$, $z = -1.38$, $p = .17$, at 50th percentile $\beta = 0.42$, $SE = 0.33$, $z = 1.27$, $p = .20$).

In summary, the results from Study 1 reveal that, under some circumstances and for some age groups, there is an effect of gestural input on children’s comprehension of a complex syntactic construction. More specifically, during the training phase when children saw videos of puppets performing an action and then heard E describe the scene with or without accompanying gesture, the older four-year-olds benefitted from seeing gesture. However, more importantly, the older four-year-olds who had received gestural input in the training and comprehension 1 tasks went on to perform significantly better in their comprehension of OCC sentences in the comprehension 2 task, when none of the children received any gestural input.

Study 2

In Study 1, the gestural scaffolding of the object cleft construction comprised two different gestural components, the referent-indicating entity gestures, and the sweeping gestures indicating the direction of action. However, it is unclear which component of the gesture was important in facilitating comprehension of the OCC for the oldest children, and thus what the underlying mechanism might be.

One possibility is that, although the gestures themselves (both entity and direction gestures) were low in iconicity (since they do not inform about the particular type of agents, or the particular types of actions, occurring in the event), the overall gestural framing of the event is more transparent when it includes both entity *and* action gestures than when it consists of entity gestures only. Specifically, the association made by having the same hand depicting both the entity gesture representing the agent (entity gesture 2), and the sweeping gesture representing the action carried out by that agent, may be necessary to enhance comprehension. However, such an enhancement would depend on children making the connection between the two successive gestures that are different in form (i.e., the first as representing the agent, the second as representing their action) in order to interpret the gestural sequence.

An alternative possibility is that the entity gestures serve to draw attention to the relative timing of the two participant roles in the larger linguistic structure, and this alone may be sufficient to highlight the unusual mention of both patient and agent before the verb.

If entity gestures help children to attend to the labelling of the second noun which, in an OCC, occurs immediately before the verb, this may help them to associate this entity with the role of pre-verbal agent as is typical of many other early learned syntactic structures such as the simple transitive. In this case, enhanced performance could occur in the absence of a more sophisticated understanding of the relation between the entity and sweeping gestures, and, indeed, possibly in the absence of a specific understanding that the entity gesture ‘stands-in’ for a particular referent.

To narrow down these possibilities, in Study 2 we removed the directional gestures to test whether abstract entity gestures, indicating the existence of the two different entities alongside their verbal labels, was sufficient for scaffolding object cleft comprehension. If performance was no longer enhanced by gestural input, this would indicate that it is something specific to the combination of entity and directional gestures that accounts for our results, suggesting that children have an understanding that the gestures represent the referents in the sentence they hear. On the other hand, if performance continued to be enhanced by entity gestures alone, this would suggest that gestures serve to highlight the connection between the second named noun and the agent of the action, by drawing attention to the mention of the second noun immediately prior to the verb. This could, but would not necessarily, indicate that children have an understanding that the entity gestures represent the referents.

One could argue that, the directional gestures themselves rather than the combination of entity and directional gestures or indeed the entity gestures alone, may have brought about the enhancement we observed in Study 1 and that, therefore, an alternative would be to eliminate the entity gestures while retaining the directional gestures. However, entity and directional gestures differ in the degree to which they are context dependent. While entity gestures glean their meaning from their temporal co-occurrence with the spoken references to the respective entities, they can stand on their own - that is, even in the absence of the directional gesture, they retain their meaning. The directional gestures we used, however, glean their meaning from the temporal and locational embedding into the gestural sequence; once entity one and two have been associated with particular locations in gesture space, the directional movement from one of these locations towards the other is what identifies one of these entities as agent and the other as patient. In the absence of the entity gestures, the directional gesture loses its meaning (and note that in our study the direction of this gestural movement could not have been informative by simply mapping its direction onto the direction observed in the video, see Method section). Although one could argue that the directional gesture could serve as an attention grabber, it is not clear why having one’s attention drawn to the utterance-final verb would particularly help in the comprehension of the preceding OCC. For this reason, we opted to remove the entity rather than the directional gestures. And for the same reason, we argue that if we observe gestural enhancement in Study 1 (entity+directional gestures) but not in Study 2 (entity gestures only), this would be attributable to the special combination of entity with directional gestures, rather than to the directional gesture alone.

Method

Participants. 15 monolingual English-speaking children (5 girls) participated in the study. The sample size was determined by establishing the number of children in the gesture condition in Study 1 who were at or above the 75th centile for age (from age 4;6), as these were the children who showed a facilitative effect of gestural input on comprehension in Study 1. The children in Study 2 were aged between 4;6 – 4;8 (*M* = 4;7). They were recruited from nurseries and preschools in the Manchester area of the UK, and through the Child Study Centre at the University of Manchester. Children at or above the 75th centile for age in Study

1 were also included in Study 2 for comparison purposes. There were 15 children from Study 1 in the gesture condition (M_{age} = 4;7, 12 girls), and 14 children from Study 1 in the no gesture condition (M_{age} = 4;7, 8 girls).

Materials and Stimuli. The same materials were used as in Study 1 with one important change. The gestures produced to co-occur with the object cleft sentences were identical to those used in Study 1, with the exception that the sweeping action was omitted, leaving only the abstract entity gestures to coincide with mention of the patient and agent respectively, as in (2);

(2) *it was the [dog] that the [duck] scratched*

[Left-hand extends outwards at chest level, slightly cupped, with palm facing the floor and remains stationery (see Figure 1a)], then [right-hand extends outwards with palm facing the floor remains stationery (see Figure 1b)]

Thus, children’s comprehension was again assessed in four tasks – baseline, training, comprehension 1 and comprehension 2.

Results and Discussion

In order to compare performance in children who received entity + action gestures, those who received entity gestures only, and those who received no gestural input, we compared the performance of those children who participated in Study 2 with the oldest children (75th centile and above) in the gesture and no gesture conditions from Study 1 as detailed in the Participants section above. Figure 4 shows the mean number of correct responses in each task, splitting the data by input type for illustrative purposes (no gesture, entity + action gestures, entity-only gesture).

Mixed effects models were fitted to the data with participant and verb as random effects and task and gesture as predictors. The reference point was set as the baseline task with entity-only gestural input to allow us to identify any effects of entity-only gestures relative to no gesture, and to ascertain whether any observed effects mirror those observed for entity+action gestures in Study 1. From Study 1, we know that there was no difference between the gesture (entity+action), and no gesture groups in the children's performance on the baseline task, and thus that effects of gesture are due to the training the children received. The simple effects here showed that there was also no difference in performance between the Study 1 groups and the entity-only gesture group on the baseline task (entity-only gesture vs. no gesture p = .52; entity-only gesture vs. entity+action gesture p = .37), confirming that prior to receiving training, the children all performed at a similar level. However, as in Study 1, there were significant interactions between task and gesture for both the training vs. baseline, and comprehension 2 vs. baseline comparisons. When comparing the no gesture group with the entity-only gesture group, children who received no gestural input performed less well on both the training and comprehension 2 tasks than those who received entity-only gestures (no gesture*training β = -1.28, SE = 0.55, z = -2.32, p = .02; no gesture*comprehension 2 β = -1.06, SE = 0.55, z = -1.92, p = .05). In contrast, there were no interactions between task and gesture when comparing the entity+action gesture group and the entity-only gesture group. This shows that there was no difference in performance on any task relative to the baseline as a function of the nature of the gestural input (entity+action gesture*training p = .64, entity+action gesture*comprehension 1 p = .27, entity+action gesture*comprehension 2 p = .97).

Thus, our results suggest that eliminating the abstract representational gestural component indicating the direction of action (the sweeping element) did not affect children's comprehension of the OCC negatively. This is evidence that at the age of 4;7 children's ability to understand agent and patient roles in the context of OCCs benefits from abstract entity gestures co-occurring with the verbal mention of different referents. This suggests that simply highlighting the mention of a referent in preverbal position appears to be sufficient for children to assign to this entity an agent-like role in the associated action. The finding that the combined entity+action gestures did not further enhance children's performance suggests that children do not appear to benefit from the occurrence of sequentially occurring abstract and fairly bleached action and entity gestures to combine these into coherent, more informative gestural scaffolds of the event.

General Discussion

In this study we set out to establish whether viewing representational co-speech gestures would increase children's comprehension of a complex syntactic construction, the object-cleft-construction (OCC). In an OCC the canonical ordering of the agent and patient is reversed and is known to lead to difficulties in interpretation even in adults. Children do not master this construction until relatively late in development, making the OCC a good candidate for testing the potential benefits of observing gestures which represent the two participants in the event, and the direction of action between them.

Age differences in the benefits of gesture

Older four-year-olds were able to use representational gestures to improve their comprehension of the OCC (training and comprehension 2 tasks), whereas three-year-olds and younger four-year-olds were not. We therefore suggest that abstract representational co-speech gestures can have a facilitative effect on children's comprehension of complex grammatical structures. This is in line with previous findings which suggest that only from around four-years of age do children begin to use iconic gestures to assist in resolving lexically ambiguous messages (Kidd & Holler, 2009), and that similarly aged children can use gestures indicating the relative positioning of entities to understand complex semantic messages (McNeil et al., 2000). At the same time, our results from Study 2 revealed that older four-year olds benefitted solely from entity gestures in their endeavour to process the syntactic structures of the events. It is difficult to imagine that the additional directional information conveyed by the action gestures would not be helpful in processing OCCs at all (since, due to their spatial mapping with the preceding entity gesture, this gestural component helps to infer the agent). We therefore conclude that four-year olds may have had difficulties, either in understanding the relation between the sequentially performed abstract representational entity and sweeping gestures, or with combining them into a larger coherent gestural scaffold for the event that is being communicated to them, at least in the context of OCCs. This process is crucial for comprehending discourse face-to-face, where stretches of speech are typically accompanied by sequences of gestures that, through their spatial coherence, track referents throughout the discourse (Gullberg, 2006). One interesting question for future research is to see at what age children are able to master this step.

Mechanisms underlying the benefits of gesture

The gestures used in the present study were more abstract than those used in previous studies in that they bore no physical resemblance to the referents, nor did they depict the actual actions performed by these referents. Finding effects with the use of gestures low in iconicity means that the mechanism underlying the beneficial role of co-speech gestures in children’s language comprehension is unlikely to be explained on the basis of iconicity alone — at least in the context of complex grammatical constructions. Rather, the cognitive mechanism at work here appears to be one of visual attention which not only seems to facilitate the processing of those parts of the verbal utterance that the entity gestures directly coincide with (i.e., the verbal referent names), but the processing of the syntactic structure as a whole (since children achieve higher accuracy in interpreting the overall event communicated to them). One possibility is that the visual gestural markers of agent and patient accompanying the sentences in our study facilitated language processing by reducing children’s cognitive load—a role that has been attributed to representational gestures in the context of math tasks, albeit in a production context (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001). Goldin-Meadow et al. (2001) proposed various mechanisms that may underlie the cognitive load reduction during production. In the context of comprehension, one possibility is that gestures, which physically embody the things that are being talked about (such as two entities), might facilitate the mapping of entities we visually perceive in the world through which we navigate, and the lexical labels we use to refer to them. Alternatively, it could be that the gestures we used fulfil a general ‘chunking function’ by bunching the stream of lexical items they co-occur with into three essential components, an agent, a patient, and a transitive action/movement direction, thus facilitating their parsing. These possibilities remain speculations in the context of the present study but offer interesting avenues for future studies.

Our postulated function of gesture in the context of processing OCC constructions means that, potentially, other non-verbal cues (such as pictures showing monochrome dots, objects to represent entities and their movements, or any other visual attention getter) *may* function in a similar way to representational gestures. However, one reason to believe that this may not be the case is that differences have been observed between the cognitive processing of co-speech gestures and other forms of communicative visual information which is more readily understandable in the absence of speech, such as when comparing pantomimes (Willems, Özyürek, & Hagoort, 2009) or instrumental actions (Kelly, Healey, Özyürek, & Holler, 2012) with co-speech gesture. What sets pictures apart from co-speech gestures is the lack of a tight intentional coupling between the verbal and visual modalities—and again, there is evidence that this perceived intentional coupling plays a crucial role in our comprehension of multi-modal messages (Holler, Kelly, Hagoort, & Özyürek, 2012; Holler, Schubotz, Kelly, Schuetze, Hagoort, &, Özyürek, 2013; Kelly, Creigh, & Bartolotti, 2010;

Kelly, Ward, Creigh, & Bartolotti, 2007). Hence, based on the fact that non-bodily visual cues (such as pictures) neither bear the inextricably close link with speech, nor the tight connection with perceived communicative intentions as co-speech gestures do, we argue that it is unlikely that the same pattern of results would be obtained with these kinds of cues. The same holds for manual movements that are performed while speaking but do not serve a communicative function, such as scratching oneself or fixing one's hair. To obtain the gestural enhancement effect we observed in the present study, we assume that more than random hand waving is needed, first and foremost a precise temporal coupling between manual movement and speech, as well as perceived intentionality in the pairing of the visual and verbal components. However, this is an important empirical and theoretical question, and one that certainly requires future exploration.

One question that remains, of course, is how using representational gestures of a more iconic nature may change the picture. To this end, we have to remain speculative, but it is possible that with more iconic gestures we may find beneficial effects on syntactic processing for even younger children, and the bleached nature of entity gestures may be one explanation for why we did not find effects of gesture for our three-year olds. We opted for this particular gesture type because, firstly, the role of gestures of a more abstract nature in children's language comprehension is considerably less well understood than that of typical iconic gestures. Secondly, while adults commonly distinguish different characters in an event using entity gestures, highly iconic gestures are less likely to be used in the context of OCC constructions, or similar sentences conveying transitive events (such as '*The frog kicked the cow'*). One main reason for this is that it is difficult to iconically depict specific characters unless one refers to their height, size or shape, or highly salient and specific features (such as whiskers); and this only works when those features are particularly characteristic for the particular referent (e.g. a tall person) and not shared with other referents in the event (e.g., a short person). Depicting referents such as *frogs* and *cows* is incredibly difficult using gestures high in iconicity (at least without entering the domain of charades). Much more common is the depicting of actions carried out by the referents (such as kicking). However, in most contexts this is unlikely to help children differentiate between agents and patients.

Finally, we consider why our three-year olds did not benefit from the representational gestures in comprehending OCCs. One reason for this may be the low iconicity of our entity gestures. However, another explanation is that three-year olds and younger four-year-olds may simply be less exposed to entity gestures in their interactions with caregivers (or when being overhearers of interactions), but that, after considerably more training, these gestures may benefit them in comprehending OCCs just as much as for the older four-year olds. A further possibility is that at the age of four, children's pragmatic abilities may be sufficiently advanced to grasp the bleached nature of entity gestures, by analogy to the referential role played by more iconic forms. Research on adults has shown that common ground leads to gestural ellipsis and reduced gestural forms (Gerwing & Bavelas, 2004; Holler & Stevens, 2007), and gestures reduced in iconicity (as in the case of our bleached entity and generic direction gestures), may be processed as such - especially in the context of *repeated exposures* to the same gestures - and more easily so with increasing age. Yet another explanation is that older four-year olds' more advanced pragmatic skills and understanding of communicative intentions account for them perceiving speech and gesture more readily as intentionally coupled than three-year olds. This latter explanation seems less likely considering children below two-years of age are already able to understand the communicative intentions associated with pointing gestures in the absence of speech (e.g. Behne et al., 2005; Gräfenhain, et al., 2009) but, of course, perceiving speech and gesture as intentionally coupled is a more complex task, and that this task is even more complex in the case of more abstract gestures. Future research is needed to tease these different explanatory contenders apart.

Longer term effects of gesture

An important finding revealed by our investigations is that older four-year-olds (4;7) were able to extend the benefits of gestural input to situations in which gesture was no longer present. In particular, they showed greater comprehension of the OCC on the comprehension 2 task in comparison to the baseline, even though on the comprehension 2 task, none of the children saw any gestures. This finding suggests that gesture provides not only a scaffold for successful communication in the here-and-now, but also the basis for longer lasting cognitive change (Cook, Yip & Goldin-Meadow, 2010; Ping & Goldin-Meadow, 2008). Ping and Goldin-Meadow (2008) assessed 5- and 7-year-olds' ability to use iconic gestures produced in the absence of the referent they represent, and in a different gesture space to that previously occupied by the referent, to solve a Piagetian conservation task. They argued that children's success on the task showed that gestures carried benefit far beyond simply linking arbitrary language to concrete objects in the physical environment to improve understanding. Similarly, in the context of the present study, we would argue that gestures produced immediately after a concrete event in the training phase served to provide children with crucial information about how the participant roles are organised in an unfamiliar syntactic construction. This information extended beyond the specific events witnessed to provide a more abstract understanding of participant roles in the OCC which could then be used to interpret new instances of the construction in the absence of gesture.

One slightly puzzling finding was that the children who received gestural input failed to benefit from this on the comprehension 1 task. We can only speculate as to why this was the case. In the training task, children witnessed the event, and heard a sentence describing that event, whereas in the comprehension 1 task, children no longer witnessed the event. One possibility is that it took children a little time to realise that they had to pay particular attention to what was said/gestured, rather than relying to some extent on seeing the event, resulting in the lack of an observable advantage for those with gestural input. There is some evidence to suggest that this may have in part accounted for the results. In Study 1, fewer children in the gesture group chose the correct picture on the first trial of the comprehension 1 task than on subsequent trials, but this was not the case for children in the no gesture group, and nor was this pattern observed for the children in the gesture group on the training or comprehension 2 tasks. Including additional trials may help to determine whether this accounted for the lack of a gesture effect on this task.

Theoretical accounts of why gestures may benefit the comprehension of syntactic structures

Our data raise interesting questions for usage-based approaches to language acquisition. Although these approaches view language acquisition and the development of the adult grammar as a gradual process which emerges from more general socio-cognitive skills such as intention reading and pattern finding (Tomasello, 2003), the exact role of gestural input in this process beyond facilitating joint attention between the child and their caregiver is unclear. We assume that children can learn language in the absence of specific gestural input. Nevertheless, our findings suggest that gesture could play a facilitative role in aspects of syntax acquisition. In principle, we would expect the beneficial effect of abstract entity gestures to generalise to structures other than the OCC, but determining the precise mechanism underlying gestural effects requires further work. For instance, effects may be most salient in non-canonical sentence structures where the prototypical mapping of referents to semantic roles is reversed, potentially increasing processing demands. On the other hand, in structures such as the English passive (*e.g. The duck was chased by the pig*), the relationship between the action and the agent will only be made clear through gesture if a child understands that the hand used to indicate the action is also subsequently used to refer to the agent of that action as, unlike in the OCC, the agent does not precede the action. If gestures simply serve to draw attention to a referent, and this is only useful when that referent occurs in a prototypical position in relation to the verb, then we might find little benefit for this structure. Moreover, abstract entity gestures may only play a role in the acquisition of complex syntactic structures that are learned relatively late in development as at younger ages when children acquire simple grammatical structures they may not have the cognitive capacities to process abstract gestural forms. It is also possible that the effects of gesture differ cross-linguistically as a function of a language's canonical word order. Finally, it is an empirical question as to when caregivers start using abstract representational gestures (abstract entity and action gestures of the kind used in the present studies) in their interactions with children, and thus whether the observed benefits might be representative of a more typical learning environment.

To conclude, the current studies provide important new data on children's ability to extract information from representational co-speech gestures low in iconicity and use this to assist in the acquisition of a complex grammatical construction. From around four-and-a-half years of age, children are able to utilise gestures to glean information about the semantic roles played by participants in a causative event when described using the non-canonical and infrequent object-cleft-construction. Since grammatical structure constitutes the foundation of human language, further explorations investigating to what extent both abstract and more iconic representational gestures can facilitate the comprehension and acquisition of syntax is an important avenue for future research. Here, we have made a first small step in this direction.

Tables, Figures & Captions

Table 1: *Model coefficients for correct responses including all four tasks (reference = baseline, no gesture, centred on 50th percentile for age)*

	β	*SE*	*z*	*p*
Gesture	-0.05	0.24	-0.21	.83
Training	0.03	0.24	0.12	.90
Comprehension 1	-0.05	0.25	-0.22	.83
Comprehension 2	-0.14	0.24	-0.58	.56
Age*Task_training*Gesture	0.11	0.05	2.34	.02
Age*Task_comp2*Gesture	0.12	0.05	2.70	.01

Figure 1: *Hand gestures to accompany OCC*

(a) left hand indicates patient referent - ‘It was the dog’

(b) right hand indicates agent referent - ‘that the duck’

(c) right hand makes sweeping movement towards left hand - ‘scratched’

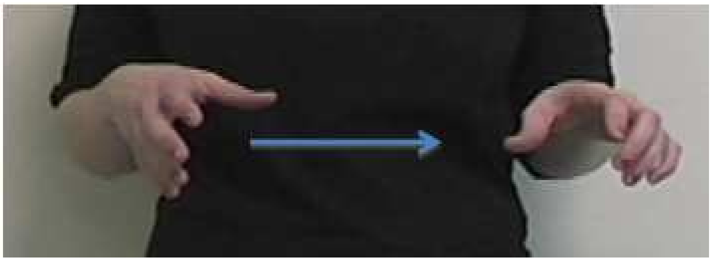


Figure 2: *Example picture cards (It was the dog that the duck scratched)*

Target Foil

Figure 3: *Study 1 mean correct responses (out of 4, and SE) in comprehension tasks.*

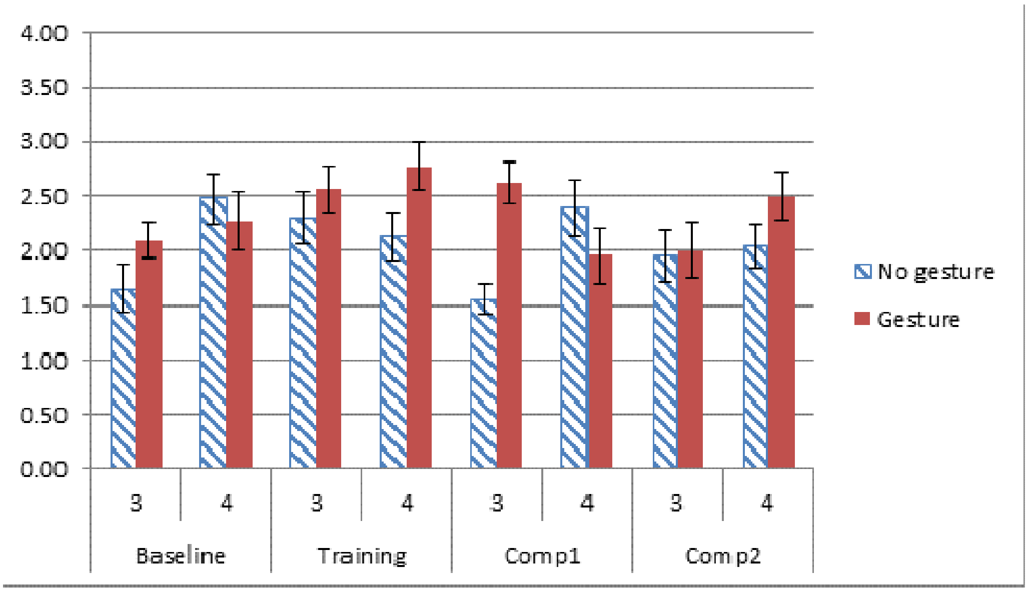
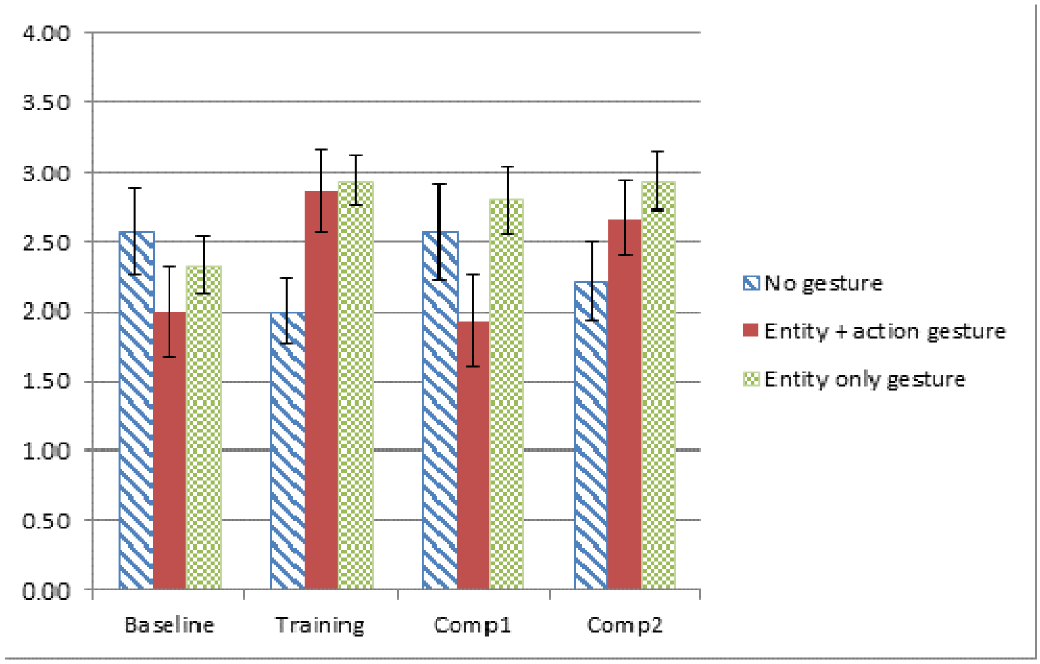


Figure 4: *Study 2 mean correct responses (out of 4, and SE) in comprehension tasks*.



Appendix: Object-cleft-constructions used in the study

It was the giraffe that the crocodile pushed.

It was the pig that the elephant hit.

It was the ladybird that the bear lifted.

It was the duck that the cow threw.

It was the fox that the dog scared.

It was the crocodile that the pig dropped.

It was the elephant that the ladybird shook.

It was the bear that the duck smelt.

It was the cow that the fox stopped.

It was the dog that the giraffe poked.

It was the giraffe that the pig squashed.

It was the crocodile that the elephant chased.

It was the pig that the ladybird carried.

It was the elephant that the bear scratched.

It was the ladybird that the duck followed.

It was the bear that the cow kicked.

It was the duck that the fox patted.

It was the cow that the dog stroked.

It was the dog that the crocodile bit.

It was the giraffe that the elephant tickled.

It was the dog that the duck scratched.

It was the cow that the ladybird tapped.

It was the crocodile that the bear hugged.

It was the fox that the giraffe punched.

References

Acredolo, L. P., & Goodwyn, S. W. (1988). Symbolic gesturing in young infants. *Child Development*, *59,* 450-466. doi: 10.2307/1130324

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, 72(2), 416-430. doi: 10.1111/1467-8624.00287

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory & Language, 44*, 169-188. doi:10.1006/jmla.2000.2752

Ambridge, B., Theakston, A., Lieven, E., & Tomasello, M. (2006). The distributed learning effect for children's acquisition of an abstract grammatical construction*. Cognitive Development*, 21, 174-93. doi:10.1016/j.cogdev.2005.09.003

Baayen, R.H., Davidson, D.J. & Bates, D.M. (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412. doi:10.1016/j.jml.2007.12.005.

Beattie G., Shovelton H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica, 123,* 1-30. doi: 10.1515/semi.1999.123.1-2.1

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. Developmental Science, 8, 492-497. doi:10.1111/j.1467-7687.2005.00440.x

Bekken, K. (1989). Is there motherese in gesture? Unpublished doctoral dissertation. University of Chicago, IL.

Bever, T. G (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language,* pp. 279-362. New York:Wiley

Brandt, S. Kidd, E., Lieven, E. & Tomasello, M. (2009). The discourse bases of relativization: An investigation of young German and English-speaking children's comprehension of relative clauses. Cognitive Linguistics, 20, 539–570. Doi: 10.1515/COGL.2009.024

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modelling study. *Journal of Child Language, 35*, 207-220. doi:10.1017/S030500090700829X

Capirci, O., Contaldo, A., Caselli, M. C., & Volterra, V. (2005). From action to language through gesture. *Gesture, 5*, 155-177. doi: 10.1075/gest.5.1.12cap

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63* (4, Serial No. 255).

Colonesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, 30, 352-366. doi:10.1016/j.dr.2010.10.001

Cook, S. W., Yip, T.K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal of Memory and Language*, 63, 465-475. doi: 10.1016/j.jml.2010.07.002

Correa, L. M. S. (1995). An alternative assessment of children's comprehension of relative clauses. *Journal of Psycholinguistic Research, 24*, 183-203. doi: 10.1007/BF02145355

Dabrowska, E., Rowland, C. & Theakston, A. (2009). Children's acquisition of questions with long distance dependencies. *Cognitive Linguistics*, *20*, 571-597. doi: 10.1515/COGL.2009.025

Ferreira, F., 2003. The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203. doi: 10.1016/S0010-0285(03)00005-7

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17,* 684–691. doi: 10.1111/j.1467-9280.2006.01767.x

Gerwing, J., & Bavelas, J.B. (2004). Linguistic influences on gesture's form. *Gesture*, 4, 157-195. doi: 10.1075/gest.4.2.04ger

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology, 91,* 720–730. doi: 10.1037/0022-0663.91.4.720

Goldin-Meadow, S., Nusbaum, H., Kelly, S., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12,* 516-522. doi: 10.1111/1467-9280.00395

Gordon, P. & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition, 36*, 227-254. doi: 10.1016/0010-0277(90)90058-R

Gräfenhain, M., Behne, T., Carpenter, M., Tomasello, M. (2009). One-year-olds' understanding of nonverbal gestures directed to a third person. *Cognitive Development, 24,* 23-33. doi: 10.1016/j.cogdev.2008.10.001

Graham, J.A. & Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. *International Journal of Psychology, 10,* 57-67. doi: 10.1080/00207597508247319

Grimminger, A., Rohlfing, K. J., & Stenneken, P. (2010). Children's lexical skills and task demands affect gestural behaviour in mothers of late-talking children and children with typical language development. *Gesture, 10,* 251-278. doi: 10.1075/gest.10.2-3.07gri

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning, 56,* 155-196. doi: 10.1111/j.0023-8333.2006.00344.x

Hupet, M. & Tilmant, B. (1989). How to make young children produce cleft sentences. *Journal of Child Language*, 16, 251-261. doi: 10.1017/S0305000900010400

Holler, J., Kelly, S., Hagoort, P., & Özyürek, A. (2012). When gestures catch the eye: The influence of gaze direction on co-speech gesture comprehension in triadic communication. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 467-472). Austin, TX: Cognitive Society.

Holler, J., Schubotz, L., Kelly, S., Schuetze, M., Hagoort, P., & Özyürek, A. (2013). Here's not looking at you, kid! Unaddressed recipients benefit from co-speech gestures when speech processing suffers. In M. Knauff, M. Pauen, I. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 2560-2565). Austin, TX: Cognitive Science Society.

Holler, J., Shovelton, H., & Beattie, G. (2009). Do iconic gestures really contribute to the semantic information communicated in face-to-face interaction? *Journal of Nonverbal Behavior, 33,* 73-88. doi: 10.1007/s10919-008-0063-9

Holler, J., & Stevens, R. (2007). An experimental investigation into the effect of common ground on how speakers use gesture and speech to represent size information in referential communication. *Journal of Language and Social Psychology*, 26, 4-27. doi: 10.1177/0261927X06296428

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137,* 297-315. doi:10.1037/a0022128

Iverson, J. M., Capirci, O., Longobardi, E., & Caselli, M. C. (1999). Gesturing in mother child interactions. *Cognitive Development, 14*, 57-75. doi: 10.1016/S0885-2014(99)80018-5

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science, 16*, 367-371. doi: 10.1111/j.0956-7976.2005.01542.x

Kelly, S. D. (2001). Broadening the units of analysis in communication: Speech and nonverbal behaviours in pragmatic comprehension. *Journal of Child Language, 28,* 325-349. doi:10.1017/S0305000901004664

Kelly, S. D., Barr, D., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language, 40,* 577-592. doi: 10.1006/jmla.1999.2634

Kelly, S. D., & Church, R. B. (1997). Children's ability to detect nonverbal behaviors from other children. *Cognition and Instruction, 15,* 107-134.

Kelly, S. D., & Church, R. B. (1998). A comparison between children's and adults' ability to detect children's representational gestures. *Child Development, 69,* 85-93. doi:10.1111/j.1467-8624.1998.tb06135.x

Kelly, S. D., Creigh, P., & Bartolotti, J. (2010). Integrating speech and iconic gestures in a Stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience, 22,* 683-694. doi:10.1162/jocn.2009.21254

Kelly, S., Healey, M., Ozyurek, A., & Holler, J. (2012). The communicative influence of gesture and action during speech comprehension: Gestures have the upper hand. Abstracts of the Acoustics 2012 Hong Kong conference published in *The Journal of the Acoustical Society of America, 131,* 3311. doi:10.1121/1.4708385

Kelly, S. D., Ward, S., Creigh, P., & Bartolotti, J. (2007). An intentional stance modulates the integration of gesture and speech during comprehension. *Brain & Language, 101,* 222-233. doi: 10.1016/j.bandl.2006.07.008

Kendon, A. (2004). *Gesture: Visible action as utterance*. New York: Cambridge University Press.

Kidd, E., & Holler, J. (2009). Children's use of gesture to resolve lexical ambiguity. *Developmental Science, 12*, 903-913. doi:10.1111/j.1467-7687.2009.00830.x

Lempert, H. & Kinsbourne, M. (1978) Children's comprehension of word order: a developmental investigation. *Child Development*, *49*, 1235-1238. doi: 10.2307/1128768

MacWhinney, B. (2000). *The CHILDES Project*. Hillsdale, NJ: Erlbaum.

Marentette, P. & Nicoladis, E. (2011). Preschoolers' interpretations of gesture: label or action associate? *Cognition, 121,* 386-99. doi: 10.1016/j.cognition.2011.08.012

McNeil, N. M., Alibali,M. W., & Evans, J. L. (2000). The role of gesture in children’s comprehension of spoken language: now they need it, now they don’t. *Journal of Nonverbal Behavior, 24,* 131-150. doi: 10.1023/A:1006657929803

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago: 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, 19,* 559-580. doi: 10.1017/S0305000900011569

Namy, L. L. (2008). Recognition of iconicity doesn't come for free. *Developmental Science,* 11(6), 841-846. doi: 10.1111/j.1467-7687.2008.00732.x

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. doi:10.1207/s15327647jcd0501_3

Namy, L. L. & Waxman, S. R. (1998). Words and gestures: infants' interpretations of different forms of symbolic reference. *Child Development* 69, 295–308. doi: 10.2307/1132165

Özçalışkan, S., & Goldin-Meadow, S. (2010). Sex differences in language first appear in gesture. *Developmental Science* 13:5, 752–760. doi: 10.1111/j.1467-7687.2009.00933.x

Özçalışkan, S., & Goldin-Meadow, S. (2009). When gesture-speech combinations do and do not index linguistic change. *Language and Cognitive Processes, 24*(2), 190-217. doi: 10.1080/01690960801956911

Özçalışkan, S., & Goldin-Meadow, S. (2011). Is there an iconic gesture spurt at 26 months? In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture*, pp.163-174. Amsterdam: John Benjamins.

Ping, R. & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology*, *44*, 1277-1287. doi: 10.1037/0012-1649.44.5.1277

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior, 5,* 172-183. doi: 10.1007/BF00986134

Rodrigo, M., Gonzalez, A., Ato, M., Rodrıguez, G., de Vega, M. & Muneton, M. (2006). Co-development of child–mother gestures over the second and the third years. *Infant and Child Development*, *15,* 1–17. doi: 10.1002/icd.412

Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science, 12*, 182-187. doi: 10.1111/j.1467-7687.2008.00764.x

Sauer, E., Levine, S.C., & Goldin-Meadow, S. (2010). Early gesture predicts language delay in children with pre- and perinatal brain lesions. *Child Development*, 81 (2), 528–539. doi: 10.1111/j.1467-8624.2009.01413.x

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teachers' gestures and speech differ. *Psychological Science*, *16*, 85-89. doi: 10.1111/j.0956-7976.2005.00786.x

So, W.-C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, *33*, 115-125. doi: 10.1111/j.1551-6709.2008.01006.x

Stanfield, C., Williamson, R., & Ozçalişkan, S. (2014). How early do children understand gesture–speech combinations with iconic gestures? *Journal of Child Language*, 41(2), 462-71. doi: 10.1017/S0305000913000019

Tomasello, M. (2003). *Constructing and language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.

Tomasello, M., Striano, T., & Rochat, P. (1999). Do young children use objects as symbols? *British Journal of Developmental Psychology, 17*, 563-584. doi: 10.1348/026151099165483

Wells, C. G. (1981). *Learning through interaction: The study of language development*. Cambridge: Cambridge University Press.

Willems, R. M., Özyürek, A., & Hagoort, P. (2009). Differential roles for pSTS and LIFG in neural integration of action and language information. *NeuroImage, 47*, 1994-2002. doi:10.1016/j.neuroimage.2009.05.066

.