

GESPIN 2019

11 - 13 September



This paper was presented at the 6th Gesture and Speech in Interaction Conference that was held at Paderborn University, Germany from September 11-13, 2019.

To cite this paper:

Mertens, U., Abramov, O., Németh, A., Kern, F., Kopp, S., & Rohlfing, K. J. (2019). Children's viewpoint: Iconic co-speech gestures and their relation to linguistic structure across two communicative genres. In: Grimmering, A. (Ed.): *Proceedings of the 6th Gesture and Speech in Interaction – GESPIN 6* (pp. 62-67). Paderborn: Universitätsbibliothek Paderborn. doi:10.17619/UNIPB/1-813

Children's viewpoint: Iconic co-speech gestures and their relation to linguistic structure across two communicative genres

*Ulrich Mertens¹, Olga Abramov², Anne Németh³, Friederike Kern³,
Stefan Kopp², and Katharina J. Rohlfing¹*

¹Paderborn University, Germany; ²CITEC, Bielefeld University, Germany;

³Bielefeld University, Germany

ulrich.mertens@uni-paderborn.de, olga.abramov@uni-bielefeld.de,
anne.nemeth@uni-bielefeld.de, friederike.kern@uni-bielefeld.de,
skopp@techfak.uni-bielefeld.de, katharina.rohlfing@uni-paderborn.de

Abstract

In this study, two different communicative genres (explanation vs. report) were elicited in 38 German preschool children at the age of 4 years. In one part of the study, explanations of a game were elicited from the child. The game involved spatial movements and figures with various geometrical shapes. In a subsequent part, children reported about a puppet and its odd behaviour to their caregiver. We examined children's viewpoint in iconic co-speech gestures and related it to the children's event structures and linguistic structures that differed in terms of transitivity. Our findings suggest that children do not use viewpoints in a unified way—which had been reported from studies with adults. In contrast, our results indicate a great variability in the ways children use viewpoint in iconic co-speech gesture. We found that different communicative genres (explanation vs. report) evoke different viewpoints in gesture, due to their different event structure and linguistic structure. During the genre “explanation”, O-VPT gestures occurred more frequently with intransitive utterances, whereas during the genre “report”, C-VPT gestures occurred more frequently with transitive utterances. Moreover, neither of the events within the communicative genres exclusively evoked one specific viewpoint.

1. Introduction

During communication, gesture provides insights into people's viewpoints. The viewpoints mainly used in iconic gesture are character viewpoint (C-VPT) and observer viewpoint (O-VPT) (McNeill, 1992). Describing an event with C-VPT gestures, a speaker focuses on the manner of the observed action and takes the first-person perspective of the character (McNeill, 1992). Research suggests that by taking the C-VPT, people conceptualize an event from an internal view and thus gain first-person knowledge (Black, Turner, & Bower, 1979; Demir, Levine, & Goldin-Meadow, 2015; Speer, Reynolds, Swallow, & Zacks, 2009). In contrast, using O-VPT within an iconic gesture, a speaker presents an event from the third-person perspective, focusing on the path of a character's action (McNeill, 1992). For example, depicting how a character climbed up a ladder, using the O-VPT, the speaker would move her or his hand from below to above, which would represent the whole character and the trajectory of the character. In contrast, taking the C-VPT, a speaker could mimic the movements of the character's hands, showing how the character actually climbed up the ladder. Why speakers take a specific perspective in gesture is uncertain and remains widely debated in the literature (e.g. Dancygier & Vandelanotte, 2017; McNeill, 1992; Parrill, 2010). McNeill (1992) hypothesizes that maximally salient or newsworthy information evokes C-VPT. He also argues that the centrality of an event and the linguistic structure (transitivity) of an utterance lead to a specific viewpoint. An event can be central or peripheral to a discourse (McNeill, 1992; Parrill, 2010; Stein & Glenn, 1975), and the structure of an event can be transitive (the verb requires a direct object), or intransitive (the verb requires no direct object). While central events of narrations and transitive utterances evoke character viewpoints, peripheral events and intransitive utterances are related to observer viewpoints in gesture (McNeill, 1992). Parrill (2010) conducted a study with adults and confirmed McNeill (1992)'s proposition that C-VPTs occurred with transitive utterances

and O-VPTs with intransitive utterances. However, she did not confirm whether the centrality of an event is related to a specific viewpoint. Furthermore, Parrill (2010) showed that the structure of an event alone is more likely to evoke a specific viewpoint than the accompanying linguistic structure of an utterance.

Events which evoked C-VPTs involved some sort of handling, use of the torso, emotions, or events which are difficult to depict from an O-VPT. Events which evoked O-VPTs included trajectories. Many event structures exclusively evoke one viewpoint in particular. However, Parrill (2010) argued, that C-VPT events like reading a newspaper mostly occur with transitive utterances (“he is reading a newspaper”) while O-VPT events with trajectories usually occur with a subject-verb-prepositional phrase (intransitive).

1.1. Children’s manner of gesturing

To our knowledge, how children’s viewpoint is related to event structures and linguistic structures has not yet been investigated. Children’s manner of gesturing strongly differs from adults, e.g. in terms of object presentation and speech-gesture synchronization (Butcher & Goldin-Meadow, 2000; Heilman, Rothi, & Valenstein, 1982; Overton & Jackson, 1973). Heilman and colleagues (1982) found that children between 3 – 6 years often use body parts to represent an object physically, whereas older children and adults use their hands as hands and therefore construct an imaginary object. The differences between children and adults might be related to children’s gestural system being coupled differently with communicative behavior. In fact, speech-gesture-systems take a while to develop the proper timing (Esteve-Gibert & Prieto, 2014) and meaningful synchrony (Butcher & Goldin-Meadow, 2000). Regarding initial communicative gestures in infants, Esteve-Gibert and Prieto (2014) observed that prosodic features of vocalization and related features of the gesture execution (the gestural stroke) became more closely related to each other temporally once infants began to produce their first words. However, the differences between adults and children in the use of the VPT within gestures were never considered. The study by Demir and colleagues (2015) found that not all children at the age of five used C-VPT during a retelling task. It was observed that children were more likely to tell better structured stories at a later age when they expressed a character’s viewpoint in gesture. It is thus reasonable to argue that children’s speech behavior is related to the way that they gesture viewpoints. Little is known about how children’s speech-gesture system is established and organized, especially across communicative genres (Kern, 2011; Streeck, 2009). Communicative genres differ within their discursive demands on the interactants and require different communicative solutions (Günthner & Knoblauch, 1995; Quasthoff, Heller, & Morek, 2017). Finding patterns in communicative behavior across communicative genres would confirm that particular communicative behaviors are context-specific, or a general phenomenon.

In this study, we examined how children’s viewpoint within iconic co-speech gestures is related to the structure of an event, and to the linguistic structure of an utterance. For this purpose, we compared our findings in two different communicative genres with findings on adults’ use of viewpoints from the existing literature.

2. Method

Preschool children ($n = 38$) from Germany at the age of 4 participated in our study. Together with their caregivers, the children visited the lab. The caregiver played an active role in our setting and interacted with the child during the communicative tasks of explanation and report. During both tasks, the children spoke about events where they either performed actions by themselves, or observed a character performing actions with objects. As the children were not given a time limit for their responses, we have a wide variation in the amount of utterances and gestures children produced during their performance. To make children’s responses comparable, we divided the total number of children’s use of gesture types by children’s total number of intonation phrases.

2.1. Procedure

The main idea behind the study was to offer the children the opportunity to gesture freely. In our pilot studies, we noticed that children felt less comfortable with an experimenter, resulting in the children speaking in short utterances with fewer gestures. We therefore decided to facilitate the caregiver playing an active role as an interaction partner, and thus to elicit communicative behavior

from children which is more natural and more directly reflects their abilities. In both tasks discussed here, the caregiver was not present during a phase in which the child engaged in an activity with an experimenter. After this activity, the caregiver who had been waiting outside came back into the room and initiated a conversation with the child. Note that we did not control for the caregiver's behavior. This is because we are convinced that children at this age are scaffolded within dialogue and more complex communicative genres, which serves as a first step on their way to being able to provide a complex monologic explanation on their own.

During the communicative genre 'explanation', the experimenter and a child played a jigsaw game designed by the authors. It consisted of dice and a 20 cm x 35 cm x 0.2 cm Plexiglas board that was painted to depict a landscape. This landscape reflected a city at night: a church, a house, a moon and a star (see: Figure 1). The shapes on the dice included a: triangle, quadrangle, circle, star, and moon. One side of the dice lacked a shape. Some pieces of the board were cut out, leaving a hole that could be covered by a fitting piece. During the game, the child was allowed to throw the dice six times. If they rolled a particular shape for the first time, the child 'flew' with a small figurine through the corresponding shape on the board. First, the experimenter explained the game to the child, and then they played it together. Afterwards, the caregiver who re-entered the room asked the child what he or she had been doing. After hearing that the child had played a game, the caregiver asked the child to explain the game in order to play it with him or her later. Our analyses only refer to the situation in which the child explains the game. During the communicative genre report, the experimenter acted out a puppet character (a dog) performing some incorrect actions. The puppet dog told the child it was excited to show them how it had learned some everyday actions from humans. For example, the dog showed the child how to eat with a spoon (but held the spoon on the wrong side) or how to drink from a bottle (but drank from the bottom of the bottle) (see: Figure 1). After performing an action, the child was allowed to correct the puppet. After the activity, and after re-entering the room, the caregiver asked the child about what he or she experienced in order to elicit a report by the child. Our analyses refer only to the situation in which the child reported the event to the caregiver.

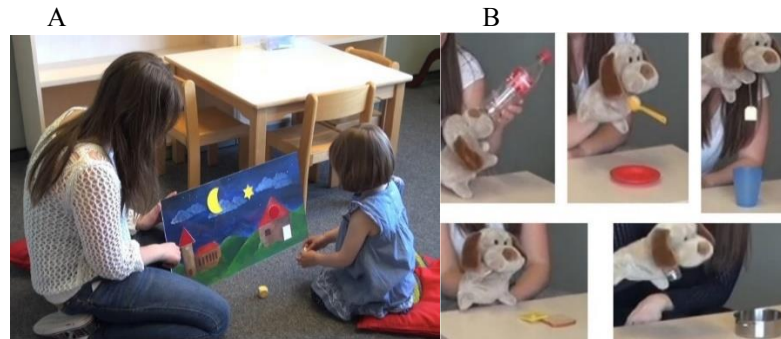


Figure 1. (A) Child plays a self-made jigsaw-puzzle-game with the experimenter (Explanation). (B) Activities the dog performed incorrectly (Report).

2.2. Coding

All verbal productions produced by narrators were transcribed and separated into intonation phrases (Barth-Weingarten, 2016; Halliday, 2015). Each utterance that was accompanied by an iconic gesture was coded regarding its transitivity (transitive: "the dog drinks water", intransitive: "the figure flies"). All iconic gestures produced by children were coded for viewpoints. Viewpoints were sorted into four categories: character viewpoint (C-VPT); observer viewpoint (O-VPT), dual viewpoint (D-VPT) and no viewpoint (N-VPT). As almost no D-VPT and N-VPT were carried out, these categories were excluded from the analyses. Children's verbal utterances were transcribed into three types of linguistic structure. The linguistic structures were transitive, intransitive and neither. We focused on transitive and intransitive utterances. We measured the agreement between the coders using Cohen's Kappa (Cohen, 1960) for speech and gesture separately. The mean Kappa value for viewpoints is $k = .860$ ($SD = .093$) and for linguistic structure $k = .887$ ($SD = .027$). During the game explanation, the events were 'throwing a dice' and 'performing flying actions with a figure'. Within the genre report, the events were: 'eating properly with a spoon', 'drinking properly from a bottle',

'placing cheese properly on a slice of bread', 'placing a teabag properly into a cup', 'putting salt with a salt shaker properly into a pot'.

3. Results

3.1. Viewpoints with respect to the event structure

During the genre explanation, children used descriptively more O-VPTs ($M = .061$; $SE = .009$) than C-VPTs ($M = .038$; $SE = .010$) but this effect was not significant ($Z = -1.708$; $r = .277$; $p = .088$). In addition, neither the event of throwing a dice nor the event of flying with a figure elicited one particular viewpoint in a dominant manner. During the genre report, children used more C-VPTs ($M = .127$; $SE = .019$) than O-VPTs ($M = .027$; $SE = .006$) reaching significant effects ($Z = -4.469$; $r = .725$; $p = .001$). However, when looking at the whole sample, we found no event that was expressed with one specific viewpoint; rather, events could be expressed with different viewpoints (see Figure 2).

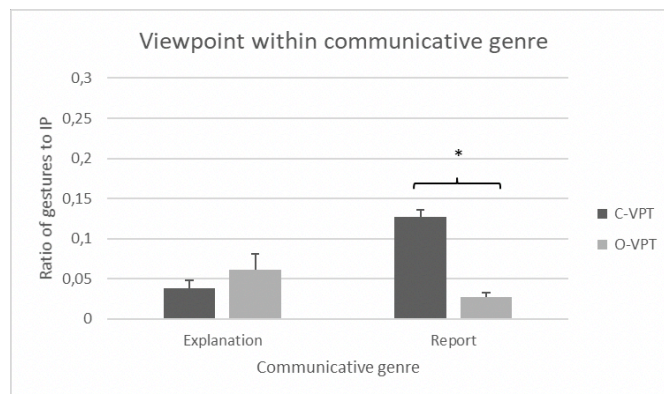


Figure 2. Ratio of gestures to intonation phrases (IP) of each viewpoint type according to communicative genres.

3.2. Linguistic structure

During the genre "explanation", children used descriptively more C-VPTs accompanied by transitive utterances ($SE = .011$; $M = .030$) than C-VPTs accompanied by intransitive utterances ($M = .014$; $SE = .005$), but this effect did not reach significance ($Z = -1.961$; $r = .319$; $p = .051$). Moreover, children used more O-VPTs accompanied by intransitive utterances ($M = .044$; $SE = .010$) than O-VPTs accompanied by transitive utterances ($SE = .005$; $M = .010$); this effect is significant ($Z = -3.043$; $r = .494$; $p = .002$).

During the genre "report", children used significantly more C-VPTs accompanied by transitive utterances ($M = .043$; $SE = .008$) than C-VPTs accompanied by intransitive utterances ($M = .011$; $SE = .003$); this effect reached significance ($Z = -3.264$; $r = .530$; $p = .001$). This means that character viewpoint was synchronized with children's utterances containing verbs with objects. Overall, children used descriptively less O-VPTs accompanied by intransitive utterances ($M = .002$; $SE = .008$) than O-VPTs accompanied transitive utterances ($M = .003$; $SE = .009$); this effect did not reach significance ($Z = -3.588$; $r = .582$; $p = .999$) (see Figure 3).

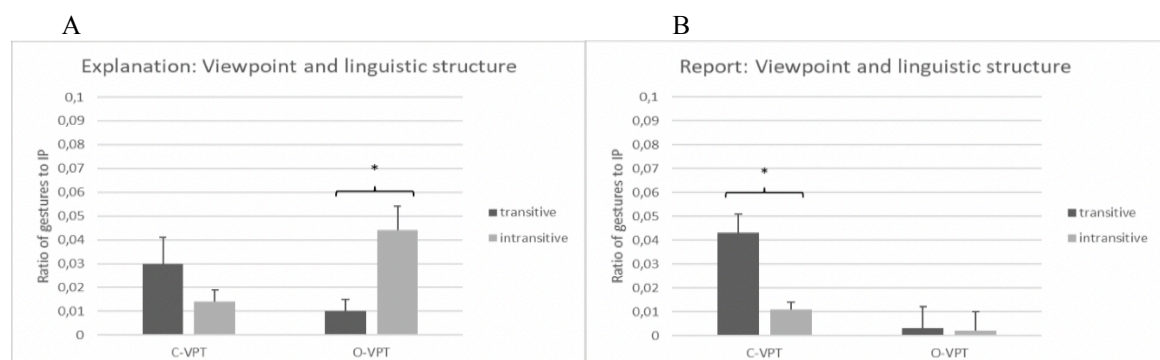


Figure 3. (A) Ratio of gestures to intonation phrases (IP) of each viewpoint type according to linguistic structure in Explanation (A) vs. Report (B).

4. Discussion

The aim of the study was to identify patterns within children's speech-gesture system. Therefore, we examined how children's linguistic structure is related to viewpoint in speech-accompanying iconic gestures across two different communicative genres with different event structures. We found that different communicative genres (explanation vs. report) evoke different viewpoints in gesture. During the genre "explanation", O-VPT gestures were more likely to occur with intransitive utterances, whereas during the genre "report", C-VPTs occurred more frequently with transitive utterances. Moreover, neither of the events evoked one specific viewpoint exclusively.

Parrill (2010) reported that adults tend to use C-VPTs when speaking about events with motoric properties, consisting of character's hands and torso, while visuo-spatial events with trajectories seem to evoke O-VPTs. In this study with adults, many events seemed to exclusively evoke one specific viewpoint in gesture (Parrill, 2010). Regarding the linguistic structure, C-VPTs seemed to occur with transitive, and O-VPTs with intransitive utterances (McNeill, 1992; Parrill, 2010). Our results extend Parrill's (2010) findings on adults' viewpoint in two ways: Firstly, we were able to analyze children's iconic co-speech gestures; secondly, we analyzed children's performance within two different communicative genres. Our findings suggest that the occurrence of children's viewpoint in gesture differs from adults in several ways. Firstly, no event was exclusively performed with one specific viewpoint in gesture. Secondly, specific viewpoints did not occur with a specific kind of linguistic structure. If Parrill's (2010) findings on adults were to be transferred to our findings, we would see some events executed exclusively with one specific viewpoint. However, in our study, the events evoked both C-VPT and O-VPT gestures. For example, using a saltshaker evoked different viewpoints: Some children represented the salt with their fingers and showed how the salt falls (O-VPT). Therefore, children represented the salt physically. Adults would be more likely mimic the holding of a salt shaker and fulfil a shaking action (C-VPT) by representing the salt shaker and the salt imaginarily (Heilman et al., 1982; Parrill, 2010). Another example is that of children showing the cube rolling over the floor, by representing the dice with their hands physically (O-VPT). There are several explanations possible for this effect. One drawn from literature suggests that children's "internal reference", or representation system, might not be fully developed (Heilman et al., 1982). Another explanation, which we favour, is that the iconic practices children use are not fully following conventionalized practices of gesturing (Streeck, 2009). Beside these differences between adults and children, our findings suggest that the events within the communicative genre of a report evoke significantly more C-VPTs than O-VPTs. This is in accordance with Parrill's (2010) findings suggesting that events with motoric properties including actions with character's hands and torso evoke C-VPTs in gesture.

For the communicative genre of an explanation, we found that intransitive utterances are more often accompanied by O-VPT gestures than transitive utterances. However, in contrast to adults, the linguistic structure (transitivity) of children's utterances did not differ when using C-VPTs. In parallel, during the communicative genre of a report, transitive utterances occurred mostly with C-VPTs. However, we found no differences between the linguistic structures that accompanied O-VPT gestures. But this could be because children used very few O-VPT gestures during the explanation phase. Therefore, we need more data to verify these results.

One explanation for why children's viewpoint does not map a specific iconic meaning to a kind of linguistic structure could be the way in which children represent objects. Children at the age of four often represent objects physically, while adults would represent the object imaginarily (Overton & Jackson, 1973), which could be due to the fact that children's iconic practices do not fully follow conventional iconic gestural practices (Streeck, 2009). This form of object representation might result in children's gestures taking an O-VPT, when adults would use a C-VPT gesture. This alternative viewpoint resulting from differences in object representation within children could be one reason why the viewpoint in gestures is not as uniform across the same linguistic structures in children as in adults.

5. Conclusion

Children's use of viewpoints in iconic co-speech gestures differs in many ways from adult usage, and between communicative genres. In addition, when comparing across different communicative genres (explanation vs. report), we do not see a unified manner in which children use their viewpoints to provide specific iconic meaning to a certain kind of linguistic structure. Instead, we found that the relation of children's viewpoint and the linguistic structure underlying children's utterances differs across communicative genres. Therefore, we can conclude that the relationship between children's viewpoints in gesture and linguistic or event structures differs from that in adults. More research is needed on how children's way of representing objects is related to their viewpoint.

Acknowledgments

This work was supported by the German Research Foundation (DFG; KE1627/3-1, KO 3510/4-1, RO 2443/8-1). We thank all the participating children and parents.

References

- Barth-Weingarten, D. (2016). *Intonation units revisited: Cesuras in talk-in-interaction. Studies in language and social interaction: Volume 29*. Amsterdam, Philadelphia: John Benjamins Publishing.
- Black, J. B., Turner, T. J., & Bower, G. H. (1979). Point of view in narrative comprehension, memory, and production. *Journal of Verbal Learning and Verbal Behavior*, 18(2), 187–198.
- Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language, culture and cognition: Vol. 2. Language and gesture* (1st ed., pp. 235–258). Cambridge: Cambridge Univ. Press.
- Dancygier, B., & Vandelanotte, L. (2017). Viewpoint phenomena in multimodal communication. *Cognitive Linguistics*, 28(3), 164.
- Demir, Ö. E., Levine, S. C., & Goldin-Meadow, S. (2015). 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.
- Esteve-Gibert, N., & Prieto, P. (2014). Infants temporally coordinate gesture-speech combinations before they produce their first words. *Speech Communication*, 57, 301–316.
- Günthner, S., & Knoblauch, H. (1995). Culturally patterned speaking practices - the analysis of communicative genres. *Pragmatics*, 5(1), 1–32.
- Halliday, M. A. K. (2015). *Intonation and grammar* (Reprint 2015). *Janua Linguarum. Series Practica: Vol. 48*. Berlin/Boston: De Gruyter; De Gruyter Mouton.
- Heilman, K. M., Rothi, L. J., & Valenstein, E. (1982). Two forms of ideomotor apraxia. *Neurology*, 32(4), 342.
- Kern. (2011). Der Erwerb kommunikativer Praktiken und Formen am Beispiel des Erzählens und Erklärens. In S. Habscheid (Ed.), *de Gruyter Lexikon. Textsorten, Handlungsmuster, Oberflächen: Linguistische Typologien der Kommunikation* (pp. 231–254). Berlin: De Gruyter.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: The University of Chicago Press.
- Overton, W. F., & Jackson, J. P. (1973). The Representation of Imagined Objects in Action Sequences: A Developmental Study. *Child Development*, 44(2), 309.
- Parrill, F. (2010). Viewpoint in speech–gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes*, 25(5), 650–668.
- Quasthoff, U., Heller, V., & Morek, M. (2017). On the sequential organization and genre-orientation of discourse units in interaction: An analytic framework. *Discourse Studies*, 19(1), 84–110.
- Speer, N. K., Reynolds, J. R., Swallow, K. M., & Zacks, J. M. (2009). Reading stories activates neural representations of visual and motor experiences. *Psychological Science*, 20(8), 989–999.
- Stein, N., & Glenn, C. G. (1975). An Analysis of Story Comprehension in Elementary School Children: A Test of a Schema. R. O. Freedle (Ed.), *New directions in discourse processing*. Norwood, NJ: Ablex.
- Streeck, J. (2009). *Gesturecraft* (Vol. 2). Amsterdam: John Benjamins Publishing.