Cognition 271 (2026) 106474

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Is a linguistic model needed to build abstract event representations?

Irene Canudas-Grabolosa [a,*], Madeline Quam [b], Marie Coppola [b], Jesse Snedeker [a], Annemarie Kocab [c]

[a] Harvard University, Department of Psychology, William James Hall, 33 Kirkland Street, Cambridge, MA 02138, USA
[b] University of Connecticut, Department of Psychological Sciences, Weston A.Bousfield Psychology Bldg, Storrs, CT 06269, USA
[c] Johns Hopkins University, Department of Cognitive Science, Krieger Hall, 3400 N. Charles Street, Baltimore, MD 21218, USA

ARTICLE INFO

Keywords:
Conceptual representation
Event representation
Relational concepts
Homesign
Two-participant events

ABSTRACT

A central question in cognitive development is whether language simply expresses pre-existing event concepts or plays a critical role in their construction and use. Recent findings from studies with infants, preschoolers and adults have raised the possibility that generic two-place relations (e.g., *cats push rabbits*) can only be represented when people have access to the transitive sentences that express them. This suggests that these concepts could be constructed as we acquire a pre-existing, external language that expresses them. To explore this hypothesis, we tested whether adult homesigners—individuals without exposure to a pre-existing language—could construct such concepts in a nonverbal imitation task. Participants viewed three instances of a given generic event (with either one or two participants), then they were given new exemplars of the same kinds (e.g., new rabbit and cat) and prompted to act. Their performance was compared to English-speaking five-year-olds. Both groups performed well in the critical two-participant condition, consistently mapping figurines of the right kind to each role. There were no group or event-type differences. Thus, homesigners have the representational resources needed to support role binding. These findings demonstrate that abstract representations of generic two-place relations can emerge without exposure to a language that models these constructions or a set of shared linguistic conventions.

1. Introduction

Is linguistic experience necessary for acquiring and using concepts? This question has a long history in cognitive science, with some proposing that language sculpts our concepts (e.g. Quine, 1960; Whorf, 1956) and thus shapes non-linguistic thought (e.g. Boroditsky, 2006), while others argue that concepts arise prior to and independent of language (e.g. Fodor, 1989; Frege, 1918). The relationship between concepts and language is likely to depend on the kind of concepts under consideration: While many object concepts might readily be acquired through direct perceptual experience (e.g., “ball”, “dog”), relational concepts (e.g., “pushing” or “same as”) are thought to pose a greater cognitive challenge. Indeed, it has been suggested that learning words that describe such relationships may help us to understand their meanings (e.g., Gentner & Rattermann, 1991; Hinzen, 2014).

Complex, compositional concepts have not been as extensively discussed in cognitive science but seem particularly amenable to a theory in which language guides conceptual construction. Take for example a generic two-place event like DOG CHASING BIRD. This concept applies to a wide range of events that vary in their perceptual features (a Labrador galloping after a robin in flight, a Chihuahua nipping at the heels of a chicken), but we can distinguish members of this category from superficially similar but relationally distinct categories (like BIRD CHASING DOG), even when the movements and entities are similar. Do we represent these generic events via a combinatorial conceptual system that exists prior to language? Or do we use language itself to build these complex representations?

Two-place asymmetrical events are of particular interest for several reasons: First, they are ubiquitous in language: 42% of the declarative sentences in child-directed speech are simple transitives (Cameron-Faulkner et al., 2003). Second, the distinction between transitive agents and patients is more reliably marked across languages than other thematic distinctions, suggesting that languages typically provide representations that could help the learner differentiate these roles (Hartmann et al., 2014). Third, two-place events require that we bind a particular kind to a specific event role. In contrast, one-participant

* Corresponding author.
E-mail address: irenecanudas@gmail.com (I. Canudas-Grabolosa).

https://doi.org/10.1016/j.cognition.2026.106474
Received 17 July 2025; Received in revised form 28 January 2026; Accepted 30 January 2026

events (dog jumping) and symmetric events (dog and cat dancing) could be categorized by merely checking to see that the right entities and actions are present. Thus, while two-argument predicates seem simple and concrete, they require a more complex semantic representation than intransitive predicates. In formal semantics (see, e.g. Dowty et al., 1980), nouns, intransitive verbs and adjectives are all modelled as sets: functions which take an argument and return a truth value (type <e, t>[1]). In contrast, two-argument predicates take an argument and return a set, which then takes another argument before returning a truth value (type <e,<e,t>>). The second argument introduces semantic recursion (functions feeding functions), arbitrariness (in the order of composition), and a semantic type with no obvious cognitive interpretation. Thus, generic two-participant relations have features that might, plausibly, place them beyond the capacity of simpler compositional systems, making them an ideal test case for evaluating whether language is necessary to form abstract relational concepts.

In this paper, we explore the representation of one- and two-participant generic events in homesigners, deaf people with no access to an existing language, in order to understand the role of external linguistic input in forming these representations. Before discussing our empirical approach, we first outline the competing theories of the relationship between two-participant events and linguistic knowledge.

1.1. Two-participant events and linguistic input

Most children are exposed to a natural language in which two-participant events can be described with transitive sentences that encode the role of each participant. These children quickly learn to interpret these sentences, and they also form stable representations of generic two-participant events. This developmental pattern is compatible with two classes of hypotheses.

1.1.1. Option 1: the conceptual capacity to represent events and their participants precedes (and guides) language development

Many theories (Carey, 2009; Fillmore, 1968; Gleitman, 1990; Pinker, 1984) assume that children come to language acquisition with stable representations of generic two-participant events, allowing them to acquire the linguistic devices that encode these events (e.g., transitive word order regularities, case marking and causative morphemes).

This perspective draws support from evidence of infants' understanding of events. For instance, six-month-olds can assign specific roles to event participants, distinguishing between scenes in which a red ball launches a green ball and those where a green ball launches a red ball (Leslie & Keeble, 1987; see Golinkoff, 1975, for similar findings in 15- to 18-month-olds). By ten months, infants differentiate which elements in a scene are relevant participants (e.g., a teddy bear is relevant when it is given to someone but not when it is merely held by someone hugging someone else; Gordon, 2003).

Further evidence comes from early verb learning. By 24 months of age, toddlers extend novel verbs to events with new objects (Waxman et al., 2009), use the number of arguments and their syntactic encoding to infer the event under discussion (Naigles, 1990; Yuan et al., 2012), and correctly interpret the syntactic cues that differentiate event participants (Gertner et al., 2006). These findings suggest that children possess an early, pre-linguistic conceptual grasp of relational events.

1.1.2. Option 2: two-participant events can only be represented with language

An alternative view (Hinzen, 2014; Shukla & de Villiers, 2021) holds that external language—particularly learning about transitive syntax—is necessary to represent generic two-participant events. Under this account, learning transitive constructions (sentences like "*the dog chases the bird*") provides the representational framework needed to separate and generalize participant roles and thus understand generic two-participant events.

Support for this hypothesis comes from studies finding that infants are less apt to make generalizations about events. For instance, 19-month-olds fail to generalize verbs to new participants unless individual identities are masked (Maguire et al., 2002), and preschoolers struggle to extend novel verbs to new agents or patients (Behrend, 1990; Imai et al., 2008; Kersten & Smith, 2002).

Shukla and de Villiers (2021) provide more direct support for this hypothesis. They used an anticipatory eye-tracking paradigm with adults and infants (12–24 months), training participants to look in anticipation to a specific two-participant event (e.g., a dog pushing a car but not a car pushing a dog). The same two participants were used for all training events. In the test trials new exemplars of the same categories were used (i.e., new dogs and cars). Only adults succeeded in anticipating for the two-participant events, while both groups anticipated for one-participant events (e.g., a dog rolling). Notably, when adults engaged in verbal shadowing they also failed with two-participant events, suggesting that disrupting access to language impairs relational event processing.

Earlier findings from de Villiers (2014) support this conclusion. In a study using an imitation task, 2- to 4-year-old children watched reversible two-participant events (e.g., a man pushing a tiger) and were then asked to imitate them. They were assigned to either a linguistic condition (e.g., "Look, the man is pushing the tiger") or a non-linguistic condition (no verbal cues). Two-year-olds performed at chance in both conditions: reversals of the event were as common as correct generalizations. Older children, however, performed much better in the linguistic condition than the nonlinguistic one. Further support comes from Hinzen et al. (2022), who found a correlation between language proficiency and success in the same imitation task among children with neurodevelopmental conditions.

There are multiple possible versions of the linguistic construction hypothesis but the strongest version predicts that without linguistic input—particularly input supporting transitive syntax—individuals will be unable to represent generic two-participant events. The current study directly tests this prediction.

1.2. Homesigners

Homesigners provide a compelling test case for evaluating the hypothesis that language is necessary for representing generic two-participant events.

Homesigners are deaf individuals who, due to a lack of accessible linguistic input, do not acquire an existing signed or spoken language. Instead, they develop their own structured communication systems (Coppola, 2020; Goldin-Meadow & Mylander, 1983). These systems are self-generated and not derived from a shared community language or from input provided by hearing caregivers (Carrigan & Coppola, 2017).

Given their lack of linguistic input, homesigners provide a unique test case for the theory that language is required to encode generic two-participant events. Under this view, we would expect homesigners to struggle with conceptualizing generic event concepts because they would not have access to the linguistic input needed to differentiate and generalize roles or support the binding of categories to those roles.

2. The current study

In this preregistered study (https://osf.io/ae23x), we adapted de Villiers' (2014) imitation task for adult homesigners. We included a group of 5-year-olds as a point of comparison: like the homesigners they have minimal formal education, but they have full access to an external language.

[1] In standard semantic type notation *e* stands for the type of entities (individuals) and *t* for truth values; ⟨e,t⟩ thus denotes functions from individuals to truth values.

Table 1
Homesigners' demographic information and summary of exposure to Lengua de Señas Nicaragüense (LSN).

ID	Demographic info		Exposure to LSN	Previous experience in testing sessions?
	Age	Gender		
1	41.3	M	He attended a program for the deaf from ages 16 to 18 (two months per year) and knows some common lexical items. He has limited contact with a few deaf individuals in the area, but none knows LSN.	Participation in research began in 1996.
2	46.2	M	He attended a vocational center for people with disabilities for six months at age 18, but had no exposure to signs. He had some deaf friends in childhood but they did not know LSN. As an adult he occasionally visited the national Deaf association in Managua but has only learned a few LSN lexical items.	Participation in research began in 1996.
3	46	F	She took weekly classes for three months in her late 30s, where she learned very few LSN signs. She has very little contact with other deaf people or signers.	Participation in research began in 2004.
4	43.6	M	In his teens, he briefly attended a deaf program with ten other students (exact duration unknown). He is friends with homesigner 5 (whom he rarely sees due to distance) and with an LSN signer who taught him some vocabulary (e.g., colors).	Participation in research began in 2023.
5	31.7	M	He has had no schooling. He has one deaf friend in town who taught him a few LSN signs and very occasionally sees his friend homesigner 4.	Participation in research began in 2023.
6	24.5	F	She attended a hearing school from ages 5 to 10 where a hearing teacher had learned a few signs. She is friends with at least one deaf woman who does not know LSN, and has started dating a deaf man who may know LSN, but thus far there is little evidence that she has acquired any LSN.	Participation in research began in 2023.
7	62.5	F	She has had no formal schooling. She lives with her sister who is deaf and much younger than her. She has learned a few LSN signs from her sister who learned LSN late in life. Her late father and late brother were also deaf, but neither of them knew LSN.	Participation in research began in 2023.
8	64.1	F	She spent three months in school at age 10. She is married to a deaf man who was part of the initial group who began developing LSN. Though she has been exposed to LSN through her husband for the past 40 years, her signing patterns substantially differ from his.	Participation in research began in 2023.

2.1. Participants

Fifteen children (Mean age = 5;4, range = 5;0–5;11, 9 females) and seven homesigners (Mean age = 46, range = 25–64, 3 females) were included in the final sample. Three additional participants were recruited but excluded due to technical failure (one child) or failure to meet the inclusion criteria (one child and one homesigner; see criteria below).

Children were recruited from the Harvard Lab for Developmental Studies database, which includes families from the greater Boston area. To be eligible, they were required to have at least 50% exposure to English and be typically developing. The study took place in the lab and families received a small toy and $10 in compensation.

Homesigners were recruited via personal contacts. To be eligible they needed to be profoundly deaf, to use signs to communicate, and to have had no access to an established sign language during childhood, including Lengua de Señas Nicaragüense (LSN). The study took place at the homesigner's house or at a community center and homesigners received $10 for their participation in a session that included several different tasks. See Table 1 for additional participant information.

The study was approved by the Institutional Review Boards (IRB) of Harvard University (children) and the University of Connecticut (adult homesigners).

3. Methods

3.1. Materials

The experiment used 26 sets of figurines (pigs, sheep, goats, cows, horses, donkeys, cats, dogs, chickens, geese, mice, rabbits, turtles, boys, girls, young women, young men, old women, old men, parrots, monkeys, crocodiles, frogs, snakes, butterflies, and lizards). Each set consisted of four distinct tokens of similar size but with different colors, postures, clothing or characteristics.

3.1.1. Procedure

Event type (one-participant vs. two-participant) was manipulated within participants. A similar procedure was used for the two groups with minimal adaptations to maximally engage each population. Children were tested by a trained undergraduate research assistant. Homesigners were tested by a trained research assistant who is also deaf. In both cases, participants were seated in front of a table with a tray placed in front of them. The experimenter sat across from them with another tray.

On every trial, the experimenter selected two sets of figurines (A and B), and demonstrated the target action three times on the tray, each time using a different token from sets A and B. The action was always performed by figurine A, while the role of figurine B varied depending on the condition: in one-participant trials, B served as a bystander, whereas in two-participant trials, B moved as a result of A's action. Figurines were assigned to particular actions but the roles of the figurines were counterbalanced across lists such that the actor for one group of participants was the bystander or patient for the other group.

After the third demonstration participants were asked to replicate what they had observed using a fourth pair of figurines. Children were given the instruction “Now it's your turn, do what you've seen”, whereas homesigners were handed the figurines and prompted to act with an open-handed forward gesture directed at them, signaling that it was their turn without indicating the figurines or modeling the action.

The study comprised 14 trials, each following the structure above, which were divided in two phases: a training phase (4 trials) and a test phase (10 trials) (Fig. 1).

The training phase began with two warm-up trials designed to introduce the task. Children were handed the figurines, and encouraged to explore and to label them to familiarize themselves (this step was omitted for homesigners). Then participants were presented with one-participant actions using only one set of figurines per trial. If they struggled to respond or made an error, the trial was repeated once. The next two trials introduced a second set of figurines. The experimenter demonstrated an event with figurine A, while figurine B remained as a bystander. As in the previous trials, feedback was provided, and if participants made an error, the trial was repeated.

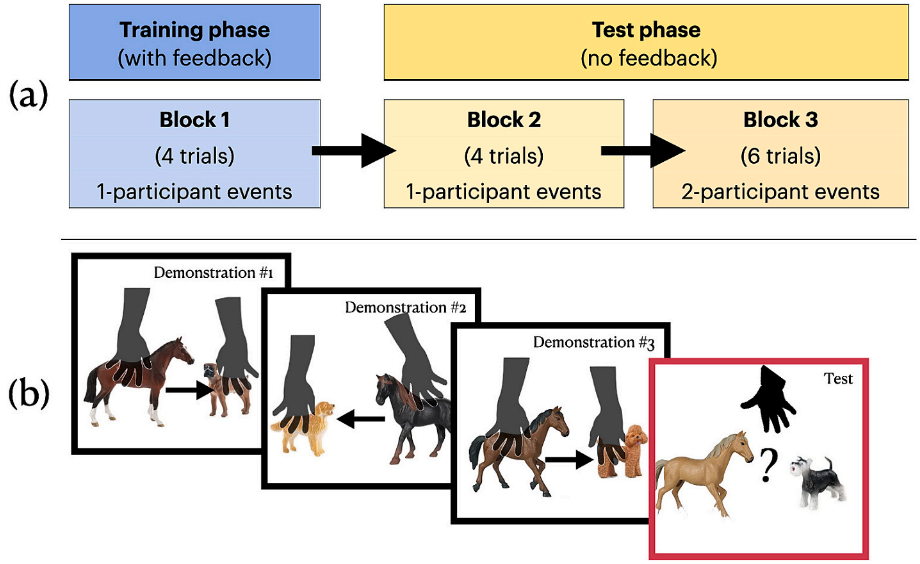

Fig. 1. (a) Procedure of the study: the study begins with a training phase where participants learned to copy one-participant events. It then proceeds to a test phase consisting of two blocks: the first one investigating one-participant events and the last one investigating two-participant events. (b) Schema of a trial: The experimenter demonstrates the action three times with different figurines and then asks or prompts the participant to copy it.

The test phase consisted of two blocks. The first block (4 trials), tested imitation of one-participant events, in which figurine A acted while B remained a bystander. The second block (8 trials) tested two-participant events, where figurine A acted upon figurine B. In this phase no differential feedback was provided, instead participants were given encouragement for any action. Children's actions were acknowledged with general praise (e.g. “Good job!”), while homesigners' actions were acknowledged with a nod and smile. Trials were not repeated, and if participants didn't respond the study continued.

The actions tested in the study were drawn from a pool of 10 one-participant actions and 10 two-participant actions. One-participant actions could be performed by a single character (jump, roll, walk, lay, run, walk backward, spin, bounce, swing, climb and descend), while two-participant actions were actions that involved one character acting

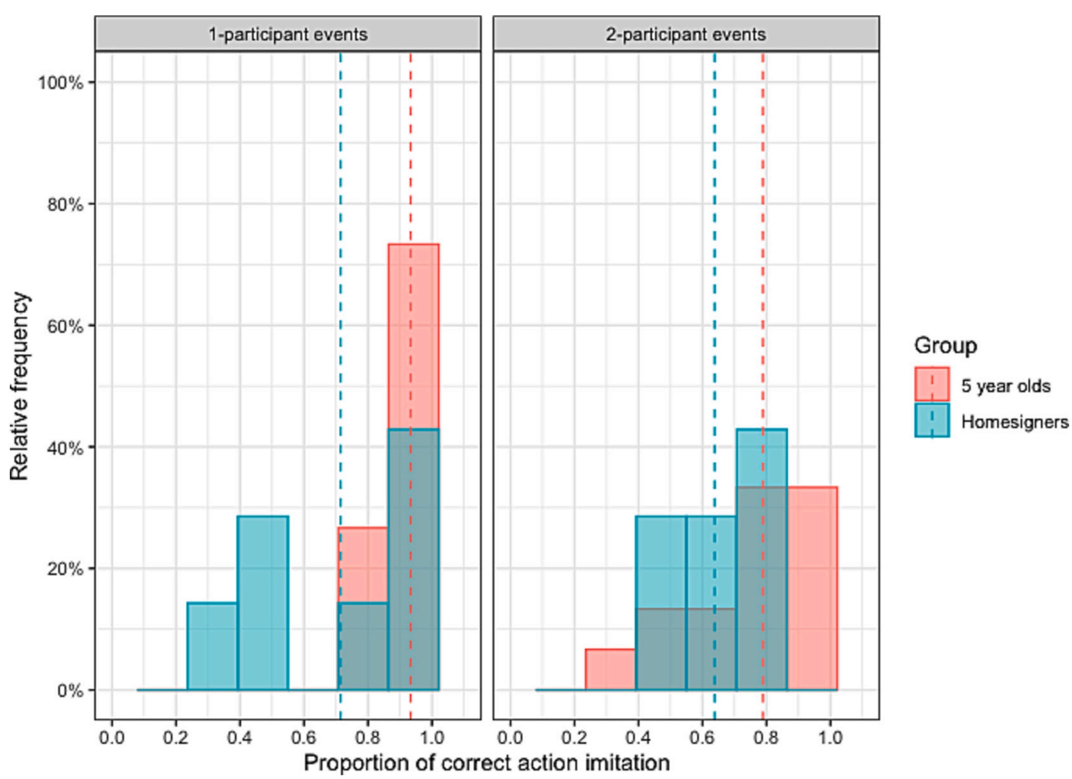

Fig. 2. Histogram of correct action imitation, divided by condition (one-participant vs. two-participant events) and group (five-year-olds in red, homesigners in blue, overlap between the groups in dark blue). The x-axis shows the proportion of correct action imitation, while the y-axis shows the percentage of participants who fall into that category. The dotted lines indicate each group's mean performance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

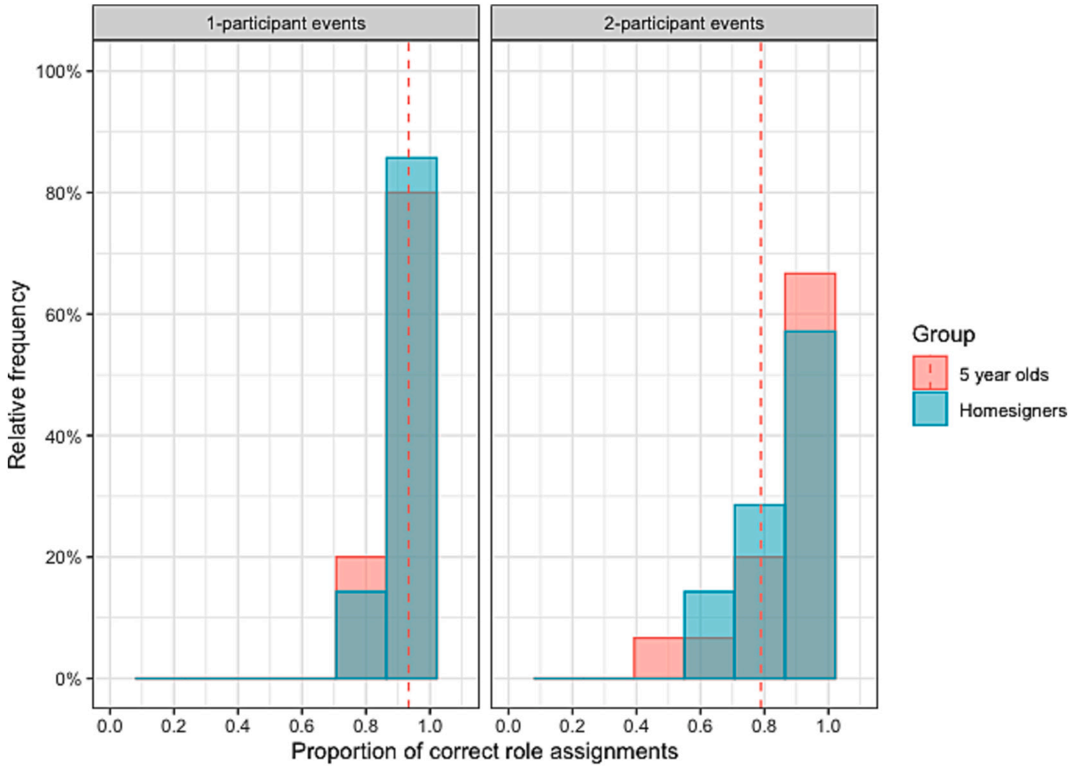

Fig. 3. Histogram of correct role assignments divided by condition (one-participant vs. two-participant events) and group (five-year-olds in red, homesigners in blue, overlap between the groups in dark blue). The x-axis shows the proportion of correct actor assignments, while the y-axis shows the percentage of participants who fall into that category. The dotted lines indicate each group's mean performance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

upon another (push, knock, touch, kick, lift, help to climb, pull, hit, drop and tap). The actions were organized in 4 different lists. Each list contained 8 one-participant actions and 6 two-participant actions, randomly drawn from the two pools above, and could be run in two possible orders, thus making a total of 8 distinct list-order combinations.

To assess children's explicit understanding of the two-participant imitation task, an additional task was conducted at the conclusion of the study. Details of this task can be found in the Supplementary Materials (https://osf.io/cyfp7).

4. Results

4.1. Coding

The study sessions were recorded and coded offline by two independent coders. To ensure that coders remained unaware of the condition and actions participants had observed, responses were isolated in individual video clips, which were then shuffled before coding.

Each trial was coded for both the action performed by the participant and the figurine that the participant used as the actor or agent in the event. Coders selected the action from a list of all actions used in the study and the actor from a list of all figurine types used in the study. The critical measure for our research question was whether the correct actor was chosen. The action was coded to assess the degree to which participants produced interpretable imitations of the actions that they saw. The inter-coder agreement was 91% for actions, and 92% for actor assignment. Any disagreements were resolved by a third coder. The experimenter's actions and actor selections were also coded following the same procedure.

4.2. Exclusion criteria

Participants were excluded if they made more than two errors in copying the actions in the first two warm-up trials, which involved only one figurine and a one-participant action. Based on this criterion, one child was excluded from the study. Additionally, individual trials were excluded from analysis if the experimenter introduced an inconsistency, such as switching the actor figurine between demonstrations. A total of nine trials (4% of the data) were excluded. One homesigner was excluded due to experimental error in 40% of their test trials.

4.3. Data analysis

To determine whether participants were engaged in imitation, we first analyzed action performance. In most cases, participants accurately reproduced the demonstrated action in a manner that was uniquely identifiable (percentage of correct action copying: one-participant trials – children: 93%, homesigners: 71%; two-participant trials – children: 79%, homesigners: 64%, see Fig. 2). Note that all 95% C.I. were well above chance, which in this case is 10% (10 possible actions to choose from).

Most errors in action imitation were small deviations in the manner in which the action was performed. When actions were grouped into broader categories (e.g., categorizing "run" and "walk" under "displacement"), the percentage of correctly imitated actions increased (one-participant trials – children: 97%, homesigners: 82%; two-participant trials – children: 87%, homesigners: 89%).

Next, we examined our main variable of interest, the accuracy of actor assignment. Accuracy in actor assignment was high across all conditions (see Fig. 3). Participants performed at near-ceiling levels when assigning the actor in one-participant events (correct assignment: children = 95%, homesigners = 96%). In our critical condition—two-

Table 2
Homesigners' individual results, divided by action copying (using the stringent criteria of individuation) and actor assignment.

ID	Demographic information		% correct action copying		% correct actor assignment	
	Age	Gender	one-participant events	two-participant events	one-participant events	two-participant events
1	41.3	M	50%	67%	75%	60%
2	46.2	M	100%	33%	100%	100%
3	46	F	50%	75%	100%	100%
4	43.6	M	100%	80%	75%	80%
5	31.7	M	100%	67%	100%	100%
6	24.5	F	75%	75%	100%	100%
7	62.5	F	25%	50%	100%	83%
8	64.1	F	This participant was excluded due to experimental error in 40% of the test trials			

participant actions—performance was also quite strong (correct assignment: children = 91%, homesigners = 89%). Separate logistic regression analyses for each group (children and homesigners) were run to confirm that the results did not differ by condition. The logistic regressions (Bates et al., 2015) included condition (one- vs. two-participant trials) as a fixed effect and participant as a random effect. To address convergence issues, item was not modelled as a random effect but instead was modelled as a fixed effect (coded as trial number). We found no significant effect of condition in either children ($\beta = -1.36$, $p = .27$) or homesigners ($\beta = -1.88$, $p = .37$).

To investigate whether the results in the critical condition–two-participant actions–were different from chance,[2] we had planned to conduct logistic regressions. These were not possible due to singular fit errors (in both samples), caused by a lack of variability in responses. Instead, the data was binarized, dividing participants into those who selected the correct actor in more than 50% of the trials and those who did not. All homesigners and 14/15 children succeeded by this criterion (in fact, 7/8 homesigners and 13/15 children made one or no errors in the task). A chi-square test confirmed that their distribution was indeed different from chance (children's $\chi^2 = 11.27$, $p < .001$; homesigner's $\chi^2 = 7$, $p = .01$; the chi-square test for homesigners simulated p-values for a Monte Carlo test with 2000 replicates, see R Core Team, 2024). Moreover, the performance of children and homesigners in two-participant trials did not significantly differ from each other as confirmed by a chi-square test ($\chi^2 = 0.49$, $p = 1$). The planned logistic regression analysis (with group and item as the main effects and participant as a random effect) could not be conducted due to singular fit errors. Table 2 presents individual homesigner's results.

5. Discussion

Our findings provide compelling evidence that the representation of two-participant events does not arise through exposure to and mastery of an external language. Homesigners, who lack access to a pre-existing language, can observe a series of two-participant events and form generalizations about the participants in those events and the roles that they play. Specifically, they can recognize that in this game butterflies push birds (and birds do not push butterflies). In fact, the homesigners were as systematic in their performance as five-year-old children who have been producing transitive sentences for three years or more. We saw no evidence that role assignment posed a particular challenge for either group: while prior studies had found that younger children had more difficulty with two-participant events than with one-participant events, both the homesigners and the five-year olds performed equally well in both conditions. These results suggest that the ability to conceptualize relational event structure–specifically, who does what to whom–is available even in the absence of exposure to a conventional linguistic system.

When we consider our findings in light of the broader literature on the origins of event representation, three questions arise.

First, what kind of conceptual representations are the homesigners (and children) using to encode the two roles in these events? Participants could succeed at our task either by encoding concrete roles tied to specific actions (e.g., the "hitter" and "hittee," see Tomasello, 1992), or by encoding more general thematic roles (such as "agent" and "patient"). Recent work has demonstrated that pre-linguistic infants are sensitive to the gestalt features that distinguish agents and patients (Papeo et al., 2024) suggesting that they may represent these abstract roles. Determining whether homesigners represent abstract thematic relations will provide strong constraints on our understanding of the origins of event representations.

Second, how can we reconcile these findings with the previous studies that have found poor representation of two-participant events in infants, children with limited language, and adults experiencing verbal interference (de Villiers, 2014; Hinzen et al., 2022; Shukla & de Villiers, 2021)? One possibility is that this reflects differences in the study populations that are independent of language experience (see below). Alternatively, the difference could lie in the nature of the tasks. The earlier studies used paradigms that required participants to make the correct generalization based on a single exemplar of the target event (sometimes repeated but always with the same participants). Our study provided participants with three different events, each enacted with unique exemplars for the agent and patient, providing clearer evidence about the scope of the rule that we were asking participants to learn (see e.g., Gómez, 2002; Xu & Tenenbaum, 2007).

Third, to what extent do these findings constrain our theories of language and event representation? Clearly these results are predicted by theories that posit that abstract event representations develop prior to and independent of language acquisition (e.g., Calvin & Bickerton, 2000; Carey, 2009; Fillmore, 1968; Gleitman, 1990; Pinker, 1984). However, they are also compatible with a weaker version of the hypothesis that language shapes event representation: successful task performance may require forming and maintaining linguistic expressions (e.g., *butterflies chase birds*) while completing the task. Homesigners possess relatively stable lexical items, unique to their individual systems, that they can combine to describe two-participant events. They may also have developed their own syntactic devices for encoding participant roles (Kocab et al., 2024; Coppola, 2002). Critically, however, whatever system they have devised does not seem to have been based on an external linguistic model: their family members—despite engaging in daily communication with them— struggle to understand the homesigners' decontextualized event descriptions (Carrigan & Coppola, 2017).

Our findings are also compatible with the possibility that the conceptualization of events in homesigners is indirectly shaped by language through the actions of the people around them and the artifacts that they encounter. While homesigners do not have access to the language of their community, they are embedded in social contexts rich in communicative interactions and symbols. It is possible that this environment provides patterns of experience that support abstraction and relational categorization. This view aligns with Gentner's structure-mapping theory (Gentner, 1988; Gentner & Rattermann, 1991), which

[2] Note that while there were 26 types of figurines that coders were choosing from, only two types of figurines were available to the participant on any trial and thus chance performance would be 50%.

emphasizes the importance of comparison and repeated exposure to relational patterns in promoting relational learning. The challenge for such a theory is to describe the facets of the environment that are needed to derive the relevant concept, how they are encoded prior to this epiphany, and the learning process that gives rise to the concept. Importantly, both of these alternatives concede the central point: that relational concepts, such as those involving asymmetrical roles between participants, can emerge without an external linguistic model.

CRediT authorship contribution statement

Irene Canudas-Grabolosa: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Madeline Quam:** Investigation, Conceptualization. **Marie Coppola:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Jesse Snedeker:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization. **Annemarie Kocab:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Acknowledgements

We are deeply grateful to the participants, the children, the homesigners, and their families, for generously giving their time and sharing their experiences. Special thanks to Marwell Guillermo Chavarría Zelaya and el Equipo Sordo (the Deaf team) in Nicaragua, whose collaboration made this research possible. We also thank the dedicated research assistants who supported data collection: Eleni C. Livingston for her work with the child participants, Kurt Gagne and Jennifer Lopez for their work with the homesigners, and Hanna S. Shine for her help with the coding.

This research was supported by the National Science Foundation under Grant No. BCS- 2116702 and BCS- 2116180. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2026.106474.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Behrend, D. A. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development, 61*(3), 681–696. https://doi.org/10.2307/1130953

Boroditsky, L. (2006). Linguistic relativity. In *Encyclopedia of cognitive science*. John Wiley & Sons, Ltd.. https://doi.org/10.1002/0470018860.s00567. s00567.

Calvin, W. H., & Bickerton, D. (2000). *Lingua ex machina: Reconciling Darwin and Chomsky with the human brain* (p. 298). The MIT Press.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science, 27*(6), 843–873. https://doi.org/10.1207/s15516709cog2706_2

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Carrigan, E. M., & Coppola, M. (2017). Successful communication does not drive language development: Evidence from adult homesign. *Cognition, 158*, 10–27. https://doi.org/10.1016/j.cognition.2016.09.012

Coppola, M. (2020). Sociolinguistic sketch: Nicaraguan Sign Language and homesign systems in Nicaragua. In O. Le Guen, J. Safar, & M. Coppola (Eds.), *Emerging sign languages of the Americas* (pp. 439–450). De Gruyter. https://doi.org/10.1515/9781501504884-015.

Coppola, M. (2002). *The emergence of the grammatical category of subject in home sign: Evidence from family-based gesture systems in Nicaragua (Unpublished doctoral dissertation)*. Rochester, NY: University of Rochester.

Dowty, D. R., Wall, R. E., & Peters, S. (1980). *Introduction to Montague semantics* (Vol. 11). Netherlands: Springer. https://doi.org/10.1007/978-94-009-9065-4

Fillmore, C. J. (1968). The case for case. In E. W. Bach, & R. T. Harms (Eds.), *Universals in linguistic theory*. Holt, Rinehart, and Winston.

Fodor, J. A. (1989). Why there still has to be a language of thought. In P. Slezak, & W. R. Albury (Eds.), *Computers, brains and minds. Australasian studies in history and philosophy of science* (7th ed., pp. 23–46). Springer. https://doi.org/10.1007/978-94-009-1181-9_2.

Frege, G. (1918). The thought: A logical inquiry. *Mind; A Quarterly Review of Psychology and Philosophy, 65*(259), 279–298. https://doi.org/10.1515/9783111546216-018

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development, 59*(1), 47–59. https://doi.org/10.2307/1130388

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman, & J. P. Byrnes (Eds.), *Perspectives on language and thought* (1st ed., pp. 225–277). Cambridge University Press. https://doi.org/10.1017/CBO9780511983689.008.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*(8), 684–691. https://doi.org/10.1111/j.1467-9280.2006.01767.x

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition: A Journal of Developmental Linguistics, 1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Noneffect of parental input on language development. *Science, 221*(4608), 372–374. https://doi.org/10.1126/science.6867713

Golinkoff, R. M. (1975). Semantic development in infants: The concepts of agent and recipient. *Merrill-Palmer Quarterly of Behavior and Development, 21*(3), 181–193.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Gordon, P. (2003). The origin of argument structure in infant event representations. In (Eds.) Smith (Ed.), *A. Brugos, l. Micciulla, & C.E.* Cascadilla press.

Hartmann, I., Haspelmath, M., & Cysouw, M. (2014). Identifying semantic role clusters and alignment types via microrole coexpression tendencies. *Studies in Language, 38*(3), 463–484. https://doi.org/10.1075/sl.38.3.02har

Hinzen, W. (2014). Recursion and truth. In T. Roeper, & M. Speas (Eds.), *Vol. 43. Recursion: Complexity in cognition* (pp. 113–137). Springer International Publishing. https://doi.org/10.1007/978-3-319-05086-7_6.

Hinzen, W., Peinado, E., Perry, S. J., Schroeder, K., & Lombardo, M. (2022). Language level predicts perceptual categorization of complex reversible events in children. *Heliyon, 8*(7), Article e09933. https://doi.org/10.1016/j.heliyon.2022.e09933

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development, 79*(4), 979–1000. https://doi.org/10.1111/j.1467-8624.2008.01171.x

Kersten, A. W., & Smith, L. B. (2002). Attention to novel objects during verb learning. *Child Development, 73*(1), 93–109. https://doi.org/10.1111/1467-8624.00394

Kocab, A., Quam, M., Coppola, M., & Snedeker, J. (2024). In *Marking event participants in Nicaraguan homesign systems [Conference presentation]* (49th Annual). Boston, MA, United States.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9

Maguire, M. J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., Slutzky, C. B., & Sootsman, J. (2002). Mapping words to actions and events: How do 18-month- olds learn a verb? In B. Skarabela, S. Fish, & A. H.-J. Do (Eds.), *Proceedings of the 26th Boston University conference of language development* (Eds.). Cascadilla Press.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357–374. https://doi.org/10.1017/S0305000900013817

Papeo, L., Vettori, S., Serraille, E., Odin, C., Rostami, F., & Hochmann, J.-R. (2024). Abstract thematic roles in infants' representation of social events. *Current Biology, 34*(18), 4294–4300.e4. https://doi.org/10.1016/j.cub.2024.07.081

Pinker, S. (1984). *Language learnability and language development (pp. xi, 435)*. Harvard University Press.

Quine, W. V. O. (1960). *Word and object*. Massachusetts Institute of Technology.

R Core Team. (2024). *R: A language and environment for statistical computing [R]*. R Foundation for Statistical Computing. https://www.R-project.org/.

Shukla, M., & de Villiers, J. (2021). The role of language in building abstract, generalized conceptual representations of one- and two-place predicates: A comparison between adults and infants. *Cognition, 213*, Article 104705. https://doi.org/10.1016/j.cognition.2021.104705

Tomasello, M. (1992). The social bases of language acquisition. *Social Development, 1*(1), 67–87. https://doi.org/10.1111/j.1467-9507.1992.tb00135.x

de Villiers, J. (2014). What kind of concepts need language? *Language Sciences, 46*, 100–114. https://doi.org/10.1016/j.langsci.2014.06.009

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty-four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology, 59*(1), 67–95. https://doi.org/10.1016/j.cogpsych.2009.02.001

Whorf, B. L. (1956). In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings from Benjamin Lee Whorf*. The MIT Press.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review, 114*(2), 245–272. https://doi.org/10.1037/0033-295X.114.2.245

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development, 83*(4), 1382–1399. https://doi.org/10.1111/j.1467-8624.2012.01783.x