Is language needed to construct two-place predicates? Event imitation in toddlers

Irene Canudas-Grabolosa (irenecanudasgrabolosa@g.harvard.edu)[1,2] & Jesse Snedeker[2]
[1]Center for Brain and Cognition, Universitat Pompeu Fabra
[2]Department of Psychology, Harvard University

Abstract

A central question in cognitive science is whether language merely expresses pre-existing concepts or provides new forms of conceptual representation. Some recent studies suggest that representing generic two-participant relations (*dogs chase birds*) requires linguistic encoding, in contrast to specific events (*this dog chased this bird*) or one-participant generic relations (*dogs jump*). We explored whether 17- to 21-month-olds, who do not yet systematically produce transitive sentences, can represent generic two-participant events. Children (n-28) completed an imitation task with either one-participant events or two-participant events. Toddlers reliably reproduced the observed action and generalized role assignments to novel exemplars (new dogs and birds), performing above chance in the two-participant condition. These results suggest that toddlers can represent transitive event relations prior to the productive mastery of transitive constructions, constraining strong claims that linguistic production is necessary for forming such representations.

Keywords: Conceptual representation; event representation; relational concepts; two-participant events

Relational concepts and language

To what extent is language necessary for representing and deploying complex conceptual structures? Does language merely label concepts that are already available, or does it play a constitutive role in forming them? These questions have a long history in cognitive science, with positions defending that language shapes and creates new conceptual representations (Quine, 1960; Whorf, 1956), while others argue that concepts arise prior and are independent of language (Fodor, 1985; Frege, 1918). The relationship between language and concepts is likely to depend on the type of concept under consideration. For example, Gentner has proposed that while many object concepts ("dog", "computer") can be readily acquired through direct perceptual experience, many relational concepts (such as "in between" or "same as") pose a greater cognitive challenge and may require linguistic scaffolding (see e,g., Gentner & Rattermann, 1991).

Complex, composed concepts have not been studied as extensively, but they seem particularly well-suited to a theory in which language guides conceptual construction. Composed concepts are those that require combining two or more constituent concepts. The present paper focuses on composed concepts that encode generic two-place events (like DOG CHASING BIRD). Concepts like these are not labels for a single event but instead denote a class of events. The class includes a wide range of perceptually-diverse situations (e.g., a husky galloping after a turkey, or a chihuahua following a sparrow), while ruling out superficially similar but conceptually distinct events (like CAT CHASING BIRD). This cannot be done simply by taking each constituent concept individually and looking for their intersection, since this procedure will fail to distinguish the intended event type from its reversal (e.g., BIRD CHASING DOG). Thus, to apply a concept of this kind, one must represent the event type (CHASING) and the two kind-concepts and link the kinds to the correct roles in the event.

Do we represent these generic two-participant events via a combinatorial conceptual system that exists prior to language? Or is language needed to build these complex representations? Concepts of this kind are a particularly interesting test case for several reasons. First, they are frequently used: approximately 42% of the declarative sentences in child- directed speech are simple transitives (Cameron-Faulkner et al., 2003). Second, languages typically differentiate transitive agents and transitive patients, via word order, case marking or agreement (Hartmann et al., 2014). This pattern raises the possibility that linguistic input could provide a model for constructing conceptual structures that support role differentiation. Or it could suggest that conceptual structures that differentiate roles are available prior to language in both evolutionary and development time (see Hartshorne & Snedeker, 2026).

Finally, while two-participant generic events are tantalizingly concrete, they require representational resources that go beyond those needed to capture other simple composed concepts. In the simplest semantic theories (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>). 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 composed concepts.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

Taken together, these considerations make generic two-participant events an informative test case for evaluating the role of language in the acquisition and use of composed concepts. Before discussing our empirical approach to that question, we briefly review the competing views on the relationship between two-participant event concepts and linguistic knowledge.

Is language needed to encode two-participant events?

Much of the literature on language acquisition assumes that infants represent events in approximately the same way as adults and that these representations guide language acquisition. For example, both syntactic and semantic bootstrapping theories assume that children can assign thematic roles to participants in events (e.g. Fillmore, 1968; Gleitman, 1990; Pinker, 1986). While this literature typically focuses on specific events (a specific dog chasing a specific bird), the abstract representations that they invoke would clearly provide a basis for generalization.

Evidence for this view comes from two sources. First there are studies with very young infants demonstrating that they can link a role to an individual in a non-linguistic task. For example, six-month-olds distinguish between scenes in which a red ball launches a green ball and those in which a green ball launches a red ball (Leslie & Keeble, 1987 ; see also Golinkoff, 1975). Second, studies of early verb learning show that prior to the productive use of transitive constructions, toddlers can make use of the linguistic structure to assign thematic roles to participants in specific events. For example, nineteen-month-olds can use the number of arguments in a sentence to infer whether an event involves one or two participants, distinguishing between transitive and intransitive verbs (Yuan et al., 2012; see also Naigles, 1990, for classic findings with 2-year-olds). By 21 months, toddlers can correctly interpret syntactic cues such as word order to differentiate event participants (Gertner et al., 2006); and by their second birthday they are able to extend novel verbs to events with new objects (Waxman et al., 2009). .

In contrast to the view that children enter language acquisition with adult-like representations of two-participant events, an alternative account holds that language itself (both as structured input and as an internalized system) is necessary for representing generic two-participant relations (de Villiers, 2014; Hinzen, 2014; Shukla & de Villiers, 2021). On this view, the kind of compositional structure required to represent such events cannot be achieved independently of language (Hinzen, 2014). Language is proposed to play a guiding role in two fundamental ways. First, learning about transitive syntax (e.g., constructions of the form "the dog chases the bird") provides the representational framework needed to separate and generalize participant roles and understand generic two-participant events. Second, once this representational framework is established, linguistic encoding is required to deploy these representations in comprehension and use, such that successful representation of two-participant events depends on access to linguistic structure.

Support for this hypothesis comes from findings suggesting that young children have difficulty linking roles not to a specific individual but to a kind. 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).

More direct evidence is provided by Shukla and de Villiers (2021), who used an anticipatory eye-tracking paradigm with adults and infants aged 12–24 months. Participants were trained to anticipate a specific two-participant event (e.g., a dog pushing a car but not a car pushing a dog) involving the same participants across trials, and were then tested on new exemplars from the same categories. Only adults showed successful anticipation for two-participant events, whereas both adults and infants succeeded for one-participant events (e.g., a dog rolling). Crucially, adults failed to anticipate two-participant events when engaged in verbal shadowing (but succeeded when engaged in non-verbal shadowing), suggesting that disrupting access to language impaired their processing of the relational event structure: when access to linguistic resources was disrupted by verbal shadowing, participants were unable to form and generalize abstract representations of two-participant events. Infants' failure, by contrast, was interpreted as reflecting insufficient linguistic experience to support representations of generic two-participant relations.

Converging evidence comes from de Villiers (2014), who investigated children's ability to represent reversible two-participant events using an imitation task. Children aged 2 to 5 years watched an experimenter enact asymmetric transitive actions (e.g., a man figurine pushing a tiger figurine) and were then asked to reproduce the event using new exemplars (e.g. a new man and a new tiger figurines). Participants were assigned to either a linguistic condition, in which the event was described verbally (e.g., “Look, the man is pushing the tiger”), or a non-linguistic condition, in which no verbal input was provided. Two-year-olds performed at chance in both conditions, producing reversals as frequently as correct reproductions. Older children, by contrast, showed markedly better performance in the linguistic condition than in the non-linguistic condition, with the oldest groups performing at ceiling in both conditions. De Villiers argued that these developmental patterns reflect children's access to linguistic encoding: younger children were unable to benefit from verbal descriptions due to immature linguistic capacities, whereas older children increasingly used linguistic input to encode and retain relational event structure, with the oldest groups not benefiting from the linguistic condition because they could already spontaneously linguistically encode such events on their own. Consistent with this interpretation, Hinzen et al. (2022) reported a correlation between language proficiency and success on the same imitation task in children with neurodevelopmental conditions.

Taken together, these findings support the claim that access to linguistic encoding facilitates the representation and retention of reversible two-participant events**.** However, the claim that exposure to language is necessary for forming such representations has been challenged by evidence suggesting that stable representations of two-participant events can emerge in the absence of conventional linguistic input.

In particular, Canudas-Grabolosa and colleagues (2026) tested homesigners —deaf adults who grew up without access to a conventional spoken or signed language (Coppola, 2020; Goldin-Meadow & Mylander, 1983)— in an imitation task parallel to the one used in de Villiers (2014). Their results show that homesigners can successfully imitate reversible transitive events despite severe limitations in early linguistic experience. These findings challenge the claim that exposure to a conventional language is required for representing two-participant event structure. At the same time, however, they do not directly address the role of linguistic encoding in these tasks. Homesigners develop their own personal structured gestural communication systems, not shared with the community they live in (Carrigan & Coppola, 2017), and they may have idiosyncratic but stable means of conveying argument structure (Kocab et al., 2024). As a result, these findings do not directly adjudicate whether access to linguistic encoding—conventional or self-generated—is necessary for representing and deploying two-participant relations.

The linguistic encoding hypothesis can be formulated in several ways. One strong version holds that the ability to produce transitive sentences is required to represent generic two-participant events. This formulation makes a clear developmental prediction: children who have not yet begun to productively use transitive constructions should fail to represent generic two-participant events while being able to represent one-participant events, a simpler representational structure. .

Our study directly targets this prediction by investigating whether toddlers below the age of two—an age at which systematic production of transitive sentences has not yet emerged (see e.g., Ninio, 1999)—can represent generic two-participant events.

Methods

In this preregistered study (https://osf.io/t86sx), we adapted the imitation task used in the previous literature (de Villiers, 2014; Canudas-Grabolosa et al., 2026) for 17 to 21 month-olds.

Participants

Twenty-eight toddlers (Mean age = 20:01 months, range 17-21, 13 females) were included in the final sample. Eight additional toddlers were recruited but excluded due to fussiness (n=1), technical failure of the equipment (n=1) or failure to pass inclusion criteria (n=6, 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.

Materials and Procedures

Toddlers were randomly assigned to one of two conditions: a one-participant event (e.g., a horse jumping while a parrot stood nearby) or a two-participant event (e.g., a horse pushing a parrot).

Materials. The experiment used 18 unique sets of action figurines, divided into three different categories: four-legged mammals (pigs, sheep, goats, cows, horses, donkeys, cats, dogs), flying animals (chickens, geese, parrots, butterflies) and people (female young, middle-aged and old; male young, middle-aged and old). Each set comprised 3 tokens of similar size but with different colors, postures, clothing or characteristics. In each trial two sets were paired with each other, making sure that toddlers would not interact with the same set twice, and that the pairs would always be cross-category (i.e.- horses with parrots, but not with dogs), to ease identification.

Procedure. Toddlers were sited on their caregiver's lab, in front of a table with the experimenter across them. Each trial started with an *exploration phase*, allowing the infant to familiarize themselves with the materials that would be used and to account for possible preferences towards an individual figurine. The experimenter presented the infant a set of three figurines (e.g., horses) and waited until the infant explored all three of them. Infants were encouraged to explore each individual figurine, both by the experimenter and the parent (e.g., by saying "Look! Have you seen this one?"). The exploration phase finished once the infant had interacted with the three figurines or clearly demonstrated they didn't want to engage anymore. Each trial used two sets of figurines, so the procedure was repeated for the second set.

Once the exploration phase was over, the experimenter proceeded to demonstrate the events. She took one figurine from each of the two sets and enacted an event while uttering an associated sound (e.g., "wee!", "thump!", related to the performed action), which was delivered with exaggerated prosody (e.g., higher pitch, elongated duration), consistent with onomatopoeic sound effects (rather than conventional lexical labels). Since toddlers trust and tend to imitate their caregivers more than they do with strangers (Shimpi et al., 2013), caregivers were also involved and then reproduced the event with new exemplars of the same kinds. Finally, the toddler was given a third pair of figurines and encouraged to act ("Now it's your turn! Let's do 'wee!'"). The experimenter waited until the child had performed an action that was recognizable or until 60 seconds had elapsed. If the toddler didn't touch the figurines or refused to imitate, demonstrations were repeated. Figurines were unique in each trial, with no set repeated between trials. 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.

.

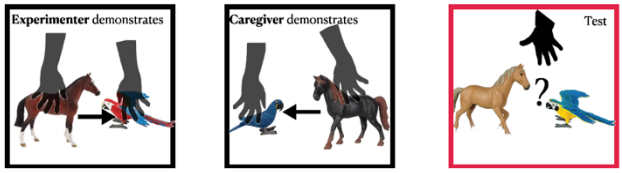

Figure 1: Structure of a test trial.

The study comprised 6 trials, each following the structure from above, which were divided in two phases: a training phase (2 trials) and a test phase (4 trials).

In the training phase, 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. In the test phase, two sets of figurines were used, and trials were not repeated.

The actions tested in the study were drawn from a pool of seven one-participant actions and six two-participant actions. One-participant actions could be performed by a single character (jump, slide, walk, sleep, walk in circles, swing, shake), while two-participant actions were actions that involved one character acting upon another (push, tap, hit, kiss, bump and sniff). The actions were organized in two different lists per condition, and could be run in two possible orders, thus making a total of 8 distinct list-order-condition combinations.

Coding Procedures

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 action was coded to assess the degree to which participants produced interpretable imitations of the actions that they saw. The critical measure for our research question was whether the correct actor (i.e., the kind of the figurine occupying the agent role in the demonstrated event) was chosen. . The inter-coder agreement was 83% for actions, and 92% for actor assignment. Disagreements were resolved by a third coded. The experimenter and parent's actions and actor selections were also coded following the same procedure.

Results

Exclusion criteria

Participants were excluded if they failed to imitate in both training trials, which involved only one figurine and a one-participant action; or if they failed to act (here defined as engaging with the figurines) in all test trials. Based on this criterion, six toddlers were excluded from the study.

Additionally, individual trials (n =12, 11% of the data) were excluded from analysis if the experimenter or the parent introduced an inconsistency, such as switching the actor figurine between demonstrations; or interfered with the task, such as pointing to the correct figurine in test trials. Finally, trials in which toddlers did not engage in any recognizable action were also excluded (n=16, 16% of the data). The final sample included 15 toddlers in the one-participant condition and 13 in the two-participant condition. On average, each toddler from the one-participant condition contributed to the data with 2.73 trials per child, while toddlers in the two-participant condition contributed with 3.08 trials per child.

Data analyses

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 identifiable (percentage of correct action copying: one-participant trials: 58%, two-participant trials: 73%, see Figure 2). Note that all 95% C.I. were well above chance, which in this case is 7.7% (13 possible actions to choose from). To investigate whether success in imitating actions differed between conditions, a Wilcoxon-Mann-Whitney test was carried out on participants' mean performance (since the planned logistic regression could not be run due to singularity effects). Results indicated no difference between conditions ($Z = -1.47$, p-value = .14).

Because 16% of the trials were excluded due to toddlers' refusal to act, we exploratorily examined whether trial exclusion differed across conditions. Specifically, we conducted a logistic regression analysis with non-response as the dependent variable, condition (one-participant vs. two-participant trials) as a fixed effect, and participant and item as random effects. The analysis revealed no evidence of a difference in exclusion rates between conditions ($\beta = 0.87$, $p = .47$; n excluded (one-participant) = 6, n excluded (two-participant) = 10).

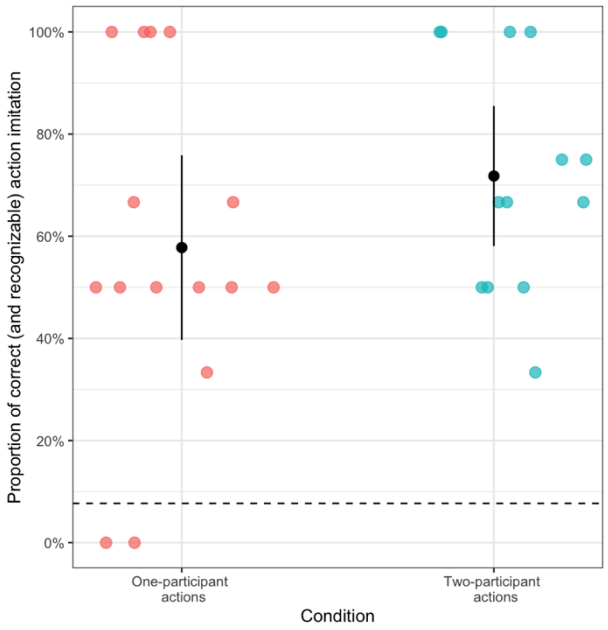

Figure 2: Proportion of accurate action imitations. Colored points indicate individual participants, while the black point marks the mean. Error bars indicate 95% confidence intervals, and the dashed line marks chance level (7.7%)

Next, we examined our main variable of interest, the accuracy of actor assignment. Accuracy was numerically above chance in both conditions (mean one-participant condition = 64%, mean two-participant condition = 72%, see figure 3).

Separate logistic regression analyses for each condition (one-participant and two-participant) were run to investigate whether results signaled above chance performance. The logistic regressions (Bates et al., 2015) included the intercept and participant as a random effect. To address convergence issues, item was not modeled as a random effect but instead was modelled as a fixed effect.

The analyses revealed that performance was slightly above chance in the two-participant condition (β=3, p=.047), but not in the one-participant condition (β=-.17, p=.79).

Although aware of the power limitations of the present sample, we also wanted to compare the two conditions. We built a logistic regression including condition (one-participant vs two-participants) as a predictor and participant as a random effect. Item, once more, was modeled as a fixed effect for the same reason. The analysis indicated that the difference between the two conditions was not reliable (β=.46, p=.46).

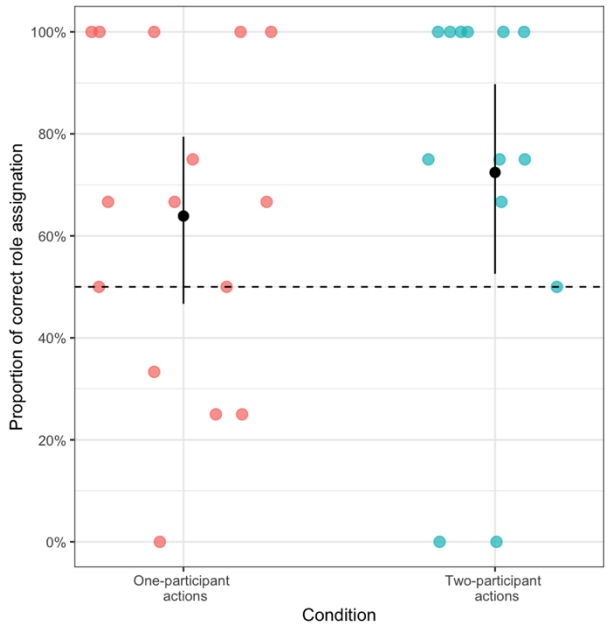

Figure 4: Proportion of accurate role assignments. Colored points indicate individual participants, while the black point marks the mean. Error bars indicate 95% confidence intervals, and the dashed line marks chance level (50%).

Discussion

In this study we provide preliminary evidence that the ability to produce transitive sentences is not required to represent generic two-participant events. Seventeen- to twenty-one-month-olds, who do not yet consistently produce utterances encoding both participants in an event (Ninio, 1999; Tomasello & Brooks, 1998), were nevertheless able to observe a series of two-participant events and form generalizations about the participants involved and the roles that they play. Specifically, toddlers were able to distinguish events in which dogs pushed birds from reversed events in which birds pushed dogs. We found no evidence that two-participant events pose a greater challenge than one-participant events; if anything, performance in the two-participant condition was slightly above chance, whereas performance in the one-participant condition was not. Together, these findings call into question the claim that generic two-participant event representations are uniquely difficult to acquire and, more specifically, that the mature production of transitive sentences is necessary for forming such representations.

One unexpected aspect of the results was the relatively poor performance in the one-participant condition. Although performance in this condition did not differ reliably from chance, the mean accuracy was numerically above (64%), raising the possibility that limited statistical power, given the small sample size, contributed to the null effect. Nevertheless, the difference with the two-participant condition, which should be interpreted with extreme caution because it was not statistically significant, was interesting. One possibility is that, in the absence of a contrasting

participant role, these events provided fewer cues for identifying the relevant dimension along which generalization was required. In the two-participant condition, role asymmetry may have highlighted the importance of role assignment in the event , whereas one-participant events may have not encouraged attention to this dimension

Could the sounds (e.g., "wee," "thump") that we used to prompt action have functioned as labels and thereby provided the linguistic support necessary to solve the task? We consider this possibility unlikely. The sounds used in the procedure were not used in transitive constructions (e.g., participants never heard things like "the dog *wees* the bird"). As such, even if these sounds were interpreted as labels (an assumption that would require independent empirical support, especially given results such as Xu et al, 2005), they would not provide the scaffolding necessary to encode role assignment, according to the linguistic encoding hypothesis. In order to do so, toddlers would need to interpret "wee" as referring not to DOG, PUSH or BIRD but to the composed concept of DOG CHASING BIRD. In other words, that would show they have the ability to represent generic two-place events.

How can the present findings be reconciled with prior studies, which report difficulties in representing two-participant events in toddlers of a similar age and in older children (de Villiers, 2014; Shukla & de Villiers, 2021)? One plausible explanation lies in the nature of the tasks. Earlier studies relied on paradigms that required generalization from multiple repetitions of a single exemplar of the target event (i.e., the same set of figurines in the imitation tasks, or the same dog and the same car in the anticipation paradigm). By contrast, the present study exposed participants to two distinct events, involving novel exemplars for both agents and patients. This design may have provided generality of the scope of the rule that we were asking participants to learn(see e.g., Xu & Tenenbaum, 2007).

If toddlers were indeed able to encode the two roles in these events, an important follow-up question concerns the nature of the representations that supported their performance. The current design does not allow us to adjudicate between different representational accounts. Infants may have succeeded by encoding relatively concrete, event specific roles tied to particular actions (e.g., the "hitter" and "hittee," see Tomasello, 1992), or by representing more abstract thematic roles such as "agent" and "patient". Recent evidence suggests that even pre-linguistic infants are sensitive to the gestalt features that differentiate agents from patients (Papeo et al., 2024), raising the possibility that abstract role representations are available early in development.

The present findings have implications for broader theories of event representation and the relationship between language and conceptual structure. Clearly, they are consistent with accounts proposing that abstract representations of two-participant events can develop prior to and independently from language acquisition (e.g., Carey, 2009; Gleitman, 1990; Pinker, 1986). At the same time, the results are also compatible with a weaker version of the linguistic encoding hypothesis.

Although the systematic production of transitive sentences does not appear to be required for representing two-participant events, it remains possible that the onset of these representations is nonetheless supported by linguistic knowledge. There is substantial evidence that language comprehension precedes production in early development, and that children below the age of two are sensitive to structural properties of transitive constructions. For instance, 19-month-olds use the number of nouns in a sentence to infer whether an event involves one or two participants, even when perceptual cues are controlled (Yuan et al., 2012), and by 21 months children can use syntactic cues such as word order to identify agents and patients in novel transitive sentences (Gertner et al., 2006). This emerging sensitivity to linguistic structure may be sufficient to support performance in the present task despite the absence of productive mastery of transitive constructions.

Note that linguistic input is not necessary to represent generic two participant events (Canudas-Grabolosa et al, 2026), as homesigners, who grew up without a conventional language, can nevertheless represent these events. However, language could play a supporting role: infants could learn these structures from their linguistic input, and thus comprehension of transitive sentences would lead to the representation of generic two-participant events. By contrast, homesigners, who do not have access to linguistic input, may have independently "discovered" generic two participant events in the process of regularizing their own outputs (which contain idiosyncratic but stable means of conveying argument structure, Kocab et al, 2024) .

Crucially, however, the present findings indicate that the ability to systematically produce linguistic descriptions encoding two-participant asymmetric relations is not necessary for conceptualizing two-participant asymmetric events. More generally, they suggest that composed concepts involving asymmetric roles between participants can emerge before children are able to describe such relations linguistically.

Acknowledgments

We are deeply grateful to the participants and their families, for generously giving their time. We also thank the dedicated research assistants who supported data collection: Eleni C. Livingston, Britney M. Balseca and Cathlyn T. Boyle.

This research was supported by the National Science Foundation under Grant No. BCS- 2116702. 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.

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

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

Canudas-Grabolosa, I., Quam, M., Coppola, M., Snedeker, J., & Kocab, A. (2026). Is a linguistic model needed to build abstract event representations? *Cognition*, *271*, 106474. https://doi.org/10.1016/j.cognition.2026.106474

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

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

Dowty, D. R., Wall, R. E., & Peters, S. (1980). *Introduction to Montague Semantics* (Vol. 11). Springer Netherlands. 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. (1985). Fodor's guide to mental representation: The intelligent auntie's vade-mecum. *Mind; a Quarterly Review of Psychology and Philosophy*, *94*(373), 76–100. JSTOR.

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., & 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.

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

Hartshorne, J. K., & Snedeker, J. (2026). *Conceptual Nativism: The origins of language are in the structure of thought*. PsyArXiv. https://doi.org/10.31234/osf.io/ ur3gj_v1

Hinzen, W. (2014). Recursion and Truth. In T. Roeper & M. Speas (Eds.), *Recursion: Complexity in Cognition* (Vol. 43, 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), 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). Marking event participants in Nicaraguan homesign systems. *Boston University Conference on Language Development 49*. 49 Boston University Conference on Language Development.

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? *Proceedings of the 26th Boston University Conference of Language Development*. Boston University Conference of Language Development.

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

Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, *26*, 619–653. https://doi.org/10.1017/S0305000999003931

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. (1986). Visual Cognition: Computational Models of Cognition and Perception. Cambridge, MA: MIT Press.

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

Shimpi, P. M., Akhtar, N., & Moore, C. (2013). Toddlers' imitative learning in interactive and observational contexts: The role of age and familiarity of the model. *Journal of Experimental Child Psychology*, *116*(2), 309–323. https://doi.org/10.1016/j.jecp.2013.06.008

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*, 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

Tomasello, M., & Brooks, P. J. (1998). Young children's earliest transitive and intransitive constructions. *Cognitive Linguistics*, *9*(4), 379–396. https://doi.org/10.1515/cogl.1998.9.4.379

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). Language, thought, and reality: Selected writings from Benjamin Lee Worf (J. B. Carroll, Ed.; print). The MIT Press.

Xu, F., Cote, M., & Baker, A. (2005). Labeling Guides Object Individuation in 12-Month-Old Infants. *Psychological Science*, *16*(5), 372–377. https://doi.org/10.1111/j.0956-7976.2005.01543.x

Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science*, *10*(3), 288–297. https://doi.org/10.1111/j.1467-7687.2007.00590.x

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