University of Pennsylvania Working Papers in Linguistics

Volume 21
Issue 1 *Proceedings from PLC 39*

Article 32

3-1-2015

Participant Structure in Event Perception: Towards the Acquisition of Implicitly 3-Place Predicates

Alexis Wellwood
alxndrw@umd.edu

Angela Xiaoxue He

Jeffrey Lidz

Alexander Williams

Participant Structure in Event Perception: Towards the Acquisition of Implicitly 3-Place Predicates

Participant structure in event perception: Towards the acquisition of implicitly 3-place predicates

Alexis Wellwood, Angela Xiaoxue He, Jeffrey Lidz, and Alexander Williams*

1 Introduction

In acquiring a semantics, children relate their experience of their world to their experience of speakers. When we study this in the lab, we often presume to understand the first part of this relation: we take for granted how the child will experience the world of our experiment, and test for how she will experience an attendant event of speech. Such presumptions are fair. But they need to be justified, when the experience we impute to the child is much richer than what the world presents objectively. In this paper we discuss one such case.

There is a hypothesis that children around the age of 2 acquire verb meanings using a heuristic that we call PAM, for "Participant to Argument Match" (Gleitman 1990, Naigles 1990, Fisher 1996, et seq.). According to PAM, the argument noun phrases in a sentence exactly match the *participant roles* in the event concept it expresses, a special subset of the roles entailed by the concept. Thus an SVO clause, with arguments S and O, is supposed to describe its event as having exactly two participant roles, the V'er and the V'ed. But the world itself is not labeled with participant roles. These are aspects of the concept under which we view a stretch of the world; and the same thing can be viewed in many different ways. Thus the hypothesis that children use PAM is falsifiable only to the extent that we are confident about how the child views events. This is the problem we mean to address.

We focus on a case that is less often discussed: transitive clauses, such as (1), that perhaps express a 3-participant concept. This case is important, since it distinguishes PAM from several other heuristics that would agree with PAM for 1- or 2-participant concepts: only PAM says that arguments match participants *for any number*.

(1) Mo stole the toy.

Imagine a child who is guided by PAM, and unfamiliar with *steal*. For her, (1) must express a concept with two participant roles (Goldberg 1995). She will therefore take any use of (1) to be about an event she can readily view in those terms, and not about one that she views under a 3-participant concept. Given this, can she understand that (1) is in fact about a theft? No, not if she is apt to view thefts as having their *victim* as a third participant, alongside the thief and his loot. But otherwise Yes: (1) is consistent with PAM so long as the role of victim is not specifically a *participant* role in her view of the theft. So we cannot say what PAM predicts for (1), or about the hypothesis that children are guided by PAM, unless we know how the child is likely to view thefts.

To that end, we report here on several experiments targeted at assessing event perception in pre-linguistic infants, following the lead of Gordon (2003). We begin by characterizing what a "participant role" is, and how PAM is meant to facilitate verb learning.

2 Participant roles and PAM

Every event of giving occurs at some time, in relation to various locations and individuals. These relations are therefore *entailed* by the event concept GIVE, which is expressed by the verb *give*. Among the entailed roles, some are special: the participant roles. These roles are psychologically distinguished or foregrounded in some way. We might think of them as the *explicit constituents* of a structured event representation associated with V, as only the mouth, eyes, and outline are ex-

* Thanks to Xuan Wang and Megan Sutton for help in preparing the experimental materials, and to Rachel Dudley for discussion of this work; Xuan Wang and Sandy Wan assisted in recruitment and coding. Thanks also to the University of Maryland's Project on Children's Language Learning lab and its lab manager, Tara Mease. This research was supported by a Social Sciences and Humanities Research Council Award (752-2010-0499) to Wellwood, and UMD Language Science fellowships to Wellwood and He.

plicit constituents of "☺", when we take this as a representation of a normal human face (Williams 2015). We can mean this as a representation of a face without denying that faces have noses.

What exactly it means to be an explicit constituent of a psychological representation is not clear, but something like this notion is presupposed in much grammatical theorizing. In figure 1, we suppose that for *give* the participant roles are just the agent, theme and recipient; that is, the giver, given, and given-to. When we represent an event with the concept GIVE, these roles are explicit. This is not to deny that all givings have locations or times, for example. These roles are entailed by the concept; they just are not separately represented.

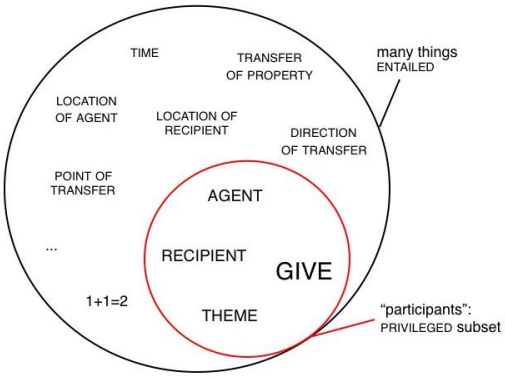

Figure 1: Participant roles are a privileged subset of the roles entailed by the GIVE concept.

Now consider what PAM can do for a learner. It says that a clause with *n* argument NPs expresses a concept with *n* participant roles. So suppose that the child hears *Anne gleebed Betty*, with two argument NPs. Guided by PAM, she will understand that the clause describes an event that is readily viewed under a 2-participant concept. If only some of the events that are salient to her can be viewed in this way, then only these can be what the sentence is about. This in turn will restrict the possible meanings for V. So suppose that only two events are salient, Anne hugging Betty, and Anne giving Betty a teddy. If the child views the hugging but not the giving under a 2-participant concept, she can make a useful inference: namely, huggings but not givings are included among the gleebings. So PAM would be useful in learning the meaning of *gleeb*.

But imagine instead that givings, while they can be viewed under a 3-participant concept, are equally likely to be viewed under an otherwise equivalent 2-participant concept, one that is satisfied only by givings, but leaves the entailed role of the thing-given implicit, roughly as the nose is implicit in our drawing of a face. In that case PAM would be no help at all; it would not decide between the hugging and the giving. The general point here deserves emphasis. PAM will not narrow down the possible meanings of a verb, if any event that can be viewed under an *n*+*m*-participant concept is equally likely to be viewed under an *n*-participant concept (Williams 2015). Therefore to test PAM with an *n*-argument clause, we must feel confident that the stimulus event in our experiment is overwhelming likely to be viewed under an *n*+*m*-participant description.

We focus our attention on cases like (1): transitive clauses whose event, in our judgment, seems likely to be viewed under a concept with more than two participant roles (Fillmore 1982, Goldberg 1995). We put (2) and (3) in the same class.

(2) Anne jimmied the box.
(3) Anne beaned Betty.

These sentences entail a third role that corresponds to no overt argument. Every stealing has a victim, in addition to a thief and some loot; every jimmying has a lever, in addition to its wielder and what the lever is used to open; and every beaning has a projectile, in addition to its thrower and the person whose head is hit with it. It also seems plausible that children will view such events under a concept that foregrounds this third role as a participant role, alongside the two others. Plausibly, they are less likely to view them under 2-participant concepts that entail the third role, but leave it at that, as for the entailed roles of time or place.

If this is correct, then PAM makes a prediction: children will fail to understand sentences like (1), (2), or (3). At best they will misunderstand them as expressing more general concepts that do not entail the third role, such as PICK UP instead of STEAL, OPEN instead of JIMMY, or HIT instead of BEAN. And this misunderstanding will persist until the counsel of PAM is abandoned. Thus we have a clear test of PAM, but only if we can first show that children view thefts, jimmyings, and beanings as we expect, under a three participant concept.

3 Assessing participant structure in event perception

At present we have no sure test for whether someone is viewing an event under a 3-participant concept, versus a 2-participant concept. But we can take steps to increase our confidence.

Suppose a subject views two scenes that differ minimally. One exhibits an event *e* to which an individual *x* bears some relation *R*—for example, an event of opening a box (*e*), in which a lever (*x*) is used as an instrument (*R*). The other scene is the same, except that *x* is inert—the lever is present but unused. Then if, according to some behavioral measure which we take to be informative, the subject does not respond to the difference as important, we may infer that she did not view the first scene under a concept which has *R* has a participant role—in our example, she viewed the first scene as an opening, and not more specifically as a jimmying (that is, as an opening-with-lever). Conversely, if she does respond to the difference, it remains possible that *R* was a participant role in her view. The hypothesis moves from possible to plausible, as we exclude other explanations of the subject's response, such as differences in the perceptual salience of *x*. If the lever is equally obvious in both our scenes, but its change in role is disregarded, then the first scene was more likely viewed as an opening than as a jimmying. Or so it is fair to argue.

Gordon (2003) makes a similar argument about event representation in pre-linguistic infants, using the habituation/switch methodology (Casasola & Cohen 2000, Werker et al. 1998). In the habituation phase, Gordon showed infants a series of videos in which one woman (Anne) hands another (Betty) a teddy bear, or in which Anne hugs Betty while holding a teddy bear. When infants reached the habituation criterion, they were presented with a new video. In the "same" condition, this was a token of the same type of scene (that is, Anne hands Betty a teddy bear, or Anne hugs Betty), but a token that they hadn't seen before. In the "switch" condition, this was a token of a different type of scene, in which the same basic action is performed, but without the teddy bear. Of interest was whether infants' attention was recaptured in response to the new videos.

Gordon found that infants dishabituated—that is, they attended more towards the new stimuli than they had at the end of the dishabituation phase—in the "switch" condition for "give" but not for "hug". He interprets this result as suggesting that infants thought something interestingly different was happening in the giving but not in the hugging. And this in turn suggests that they represented the hugging scenes *as huggings*, under a concept that does not entail the role of the teddy, and the giving scenes *as givings*, under a concept that does. Had the infants failed to notice or care about the disappearance of the teddy bear from the scenes of giving, it would seem unlikely that they were viewing that scene under a concept that entails that role.

Gordon's method shows, we believe, that the noticed role is entailed by the child's view of the event, while the unnoticed role is not. But certainly it cannot prove the stronger conclusion, that the noticed role is furthermore a participant role. Nonetheless, it does lend some credence to that conclusion if we presume, as seems reasonable, that one will not attend equally to every role entailed by the concept under which one views a thing. Human faces have noses, but we attend more to the eyes, which seem to have a distinguished role. Thefts in general have locations, but perhaps when we view an event as a theft, we are more likely to attend to the thief and the victim. Thus in our own work with infants, reported briefly in section 6, we adapt Gordon's methods to our own purposes, in designing tests of PAM.

In this paper, however, we concentrate on preparatory studies, in which we sought to find a correlative measure of the participant-structure distinction in adults. We used a similarity judgment task to assess the adult view of the stimuli we prepared for studies with infants.

4 Experiment 1: Participants match arguments

In our Exp.1, we had adults perform a similarity judgment task with stimuli like those that Gordon (2003) used for infants. In each trial we first showed subjects a video that involves a teddy bear. It either shows Anne giving Betty the teddy, or hugging Betty while holding the teddy. We then showed a second video, just like the first, except that the teddy is absent. Participants are then asked to rate "how similar these two videos were". We make no mention of "events". We recorded their explicit ratings and, importantly, also their reaction times, a more implicit measure.

Study participants. We recruited 12 adult participants from the University of Maryland undergraduate community. Participants were invited to our lab on campus, where they were asked for consent in line with the University of Maryland's IRB protocols. They were debriefed about the goals of the study afterwards. Participants received course credit or $10 for their participation.

Method. We used a Similarity Judgment Task (SJT) in which study participants were asked to judge the similarity between two videos. The structure of a single trial involved showing two videos one after another on a computer screen. Immediately following presentation of the videos, the question "How similar were these two videos" was presented on the screen, together with a 1–7 scale where '1' was marked as 'totally not similar' and 7 as 'totally similar'. Figure 2 illustrates the sequence for one trial. Participants were instructed to make judgments as quickly as possible, without sacrificing accuracy; if they were unsure, they had the option to replay the video by pressing the spacebar. There were 32 trials in total, among which 8 trials were *giving-type comparisons*—giving-with- vs. without-teddy, 8 were *hugging-type comparisons* – hugging-with vs. without-teddy, and the remaining 16 trials were *token comparisons*—different video tokens for each scene type (for example, give-with-teddy Token 1 vs. give-with-teddy Token 2). These three comparisons were within-subject conditions. Each participant completed all 32 trials in pseudo-randomized order.

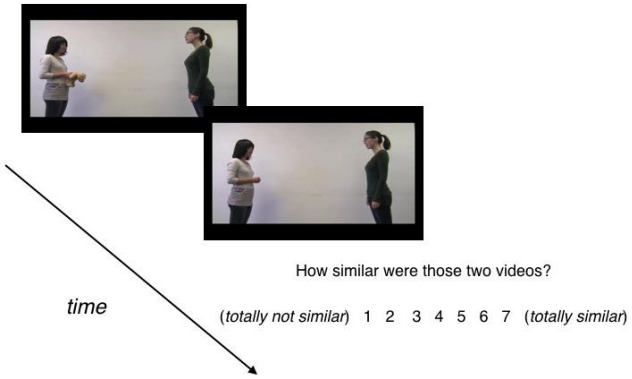

Figure 2: Schematic of the trial structure for our adult similarity judgment experiments.

Measures. Participants' rating score was our explicit measure. We also measured the time taken to make a rating decision, as an implicit measure of the processes involved in making the decision. In general, we suspected that the criteria used to make a similarity judgment are not likely to be the same as are used to encode a scene. Rating scores were measured for all trials, but reaction times (RTs) were only measured on trials without replay. The rationale for this is that reaction times after seeing the videos more than once may not be an accurate measure of the same sort of processing as seeing the videos once. In this sample of 12 participants (32*12=384 trials), the replay option was used on 51 trials in total; those trials were excluded from the RT analysis.

Predictions. Since *token comparisons* compare video tokens of the same type, the highest rating score (most similar) for this condition is expected. Apart from this, if giving-type contrasts are viewed differently from hugging-type contrasts, we expect a difference either in rating scores or in RTs, or both; if, however, the two contrasts are viewed similarly (for example, both contrasts

merely involve a difference in the teddy-bear's presence/absence), then we expect no significant difference between the two conditions in either measure.

Results. Rating scores and RTs were analyzed separately and plotted on separate graphs (Figure 3: rating scores on left, and RTs on right). In each graph, the conditions (*giving-type comparisons*, *hugging-type comparisons*, and *token comparisons*) are plotted along the x-axis, and the dependent measures on the y-axis, illustrating means across all participants and error bars showing standard errors. It is apparent that the token comparisons, not surprisingly, elicited the highest overall rating score. Two-tailed paired t-tests revealed no significant difference in rating scores between giving- (M=4.63, SD=0.90) and hugging-type (M=4,64, SD=1.05) comparisons: t(11) = -0.11, p = 0.91), but a significant difference in RTs – it took participants longer to make judgments for giving- (M=1546.03ms, SD=447.39ms) than for hugging-type (M=1151.09ms, SD=391.07ms) comparisons: t(11) = 2.0276, p = 0.07.

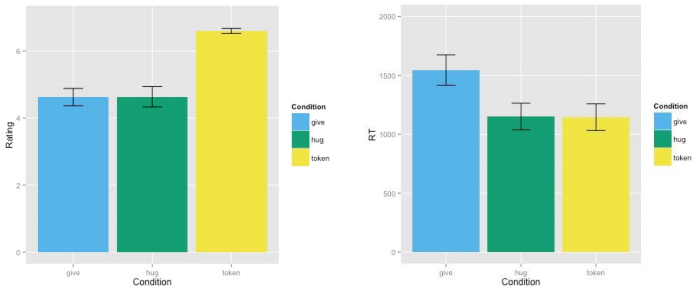

Figure 3: Mean rating scores (left) and RTs (right) by condition in Experiment 1.

Discussion. Lack of difference in the rating score suggests that the asymmetry we are interested in was not reflected in the explicit measure. This is not so surprising. There are many ways in which two things can be similar or different. And certainly, for both of our scene-types, the hugging and the giving, the gross difference between a video with and without a teddy bear is substantial, in contrast to the difference between tokens of the same type. And yet, although subjects assigned these contrasts similar rating scores, we did find an informative difference: judging by their *response times*, the similarity judgments were not made with equal ease between the two scene-types. It took subjects more time to make their similarity judgment for giving- than for hugging-type contrasts. This longer time taken in giving-type comparisons may be attributable to the surprise involved in not seeing an element that, because it fills a privileged role, is normally expected for scenes that fall under the concept GIVE. For the hugging-type comparison, however, if the teddy bear does not fill a role required for the concept HUG, then its absence should elicit no surprise.

5 Experiments 2-4: Participants may not match arguments

Prelinguistic infants and adults differentiate between giving- and hugging-type scenes with respect to the participation of a third entity. Such concepts as GIVE and HUG are typically expressed using sentences in which, plausibly, the number of syntactic arguments matches the number of participant roles. We now turn to the cases of primary interest for us, namely those in which there is a potential mismatch between syntactic and conceptual information.

In experiments 2-4, we extended our method to cases that we think people will represent under a 3-participant concept. We want stimulus scenes that can be described using sentences like those in (4), and are likely to be viewed, whether by infants or by adults, in ways that represent the roles of the lever (4a), the victim (4b), and a projectile (4c) as participant roles, despite their not being required syntactic arguments for the relevant verbs.

(4) a. Anne jimmied the box.
b. Anne stole the box.
c. Anne beaned Betty.

To this end we conducted three experiments (Exps.2-4), each with two important comparisons. *Critical comparisons* contrasted two scenes where there is, potentially and plausibly, a difference in participant structure. In one, an object is plausibly involved as a participant in an event, e.g., a lever is involved as an instrument in opening a box. In the other, the same object is present but inert. *Perceptual comparisons* are between two scenes with no such difference: the second video merely reverses the orientation of the action in the first.

Study Participants & Task (Exps.2-4). In each of Exps.2-4, we recruited 12 adult participants from the University of Maryland undergraduate community, who were either given course credit or monetary reward for their participation. Exps.2-4 used the same similarity judgment method as in Exp.1, except we eliminated the option to replay the video by pressing the spacebar. Specific differences in the stimuli are listed separately for each experiment.

Stimuli & Predictions (Exp.2). In the JIMMY/OPEN experiment, one of the videos in a pair features a girl opening the box with a big hook. In the *critical* comparison for this experiment, the orientation of the action is the same, but the girl is using her hand to open the box, while the hook is held passively in her other hand. In *perceptual* comparisons, she opens the box with the hook, but she does so from the opposite side. If the tool is represented as important to the first scene, then adults should find the critical comparison less similar than they do the perceptual comparison. As above, the similarity difference may be observed either in the explicit judgment, the reaction time, or both.

Results (Exp.2). Rating scores and RTs were analyzed separately and plotted on separate graphs (Figure 4: rating scores on left, RTs on right). Token comparisons elicited the highest rating score. Similar to Exp.1, there was no significant difference in rating scores between critical (M=3.46, SD=1.52) and perceptual (M=3.23, SD=1.22) comparisons: $t(11) = 0.96$, $p = 0.36$, but a significant difference in reaction times: participants took longer to make judgments for critical (M=1518.33ms, SD=507.13ms) than for perceptual comparisons (M=1124.76ms, SD=381.67ms): $t(11) = 2.09$, $p = 0.06$.

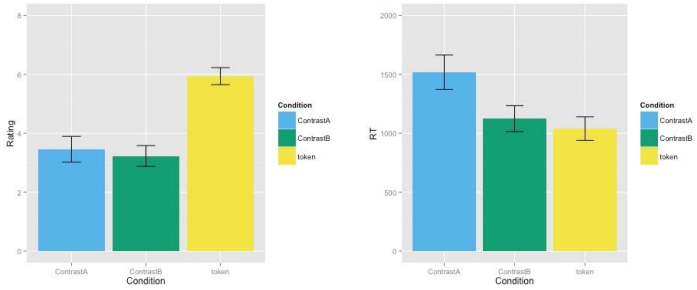

Figure 4: Mean rating scores (left) and RTs (right) by condition in Experiment 2.

Stimuli & Predictions (Exp.3). In the STEAL/PICKUP experiment, one video shows a girl picking up a toy truck from out of another girl's hands. In *critical* comparisons, the other video shows the second girl simply standing by, while the toy is picked up from the table in front of her. Here, she is just a bystander. In *perceptual* comparisons, the second girl participates in both videos, just the orientation of the action is reversed. If the second girl's role as victim is a participant role, we expect the change to bystander to be less similar than the change in orientation. As above, this difference may be evident in similarity judgments, reaction times, or both.

Results (Exp.3). Rating scores and RTs were analyzed separately and plotted on separated graphs (Figure 5: rating scores on left, and RTs on right). The token comparisons elicited the highest rating score. Like Exp.1 and Exp.2, in this experiment we still found significant difference in RTs between critical (M=1228.20ms, SD=304.49ms) and perceptual (M=838.94ms, SD=287.15ms) comparisons: $t(11) = 2.3987$, $p = 0.04$. But unlike Exps.1-2, we also found significant difference in rating scores - $t(11) = -3.2612$, $p < 0.01$: perceptual comparisons (M=5.47, SD=0.84) were judged

more similar than critical comparisons (M=4.07, SD=1.41).

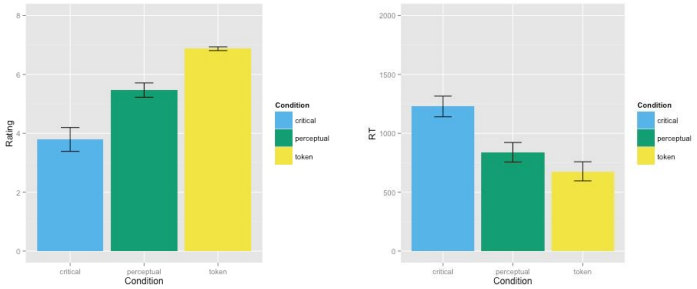

Figure 5: Mean ratings (left) and RTs (right) by condition in Experiment 3.

Stimuli & Predictions (Exp.4). In the BEAN/HIT experiment, the base case features one girl hitting another girl by throwing a blue ball. In *critical* comparisons, the first girl again hits the other, but this time the ball does not leave her hand. The ball is not acting as a projectile. In the *perceptual* comparisons, again just the orientation of the action is reversed. Here, we're interested in whether the change in the ball's role making contact as a projectile or not is seen as important to the similarity judgment. If it is, then we expect that change to be seen as less similar than the change in orientation. And again, this difference may be evident in similarity judgments, reaction times, or both.

Results (Exp.4). Rating scores and RTs were analyzed separately and plotted on separated graphs (Figure 6: rating scores on left, RTs on right). The token comparisons elicited the highest similarity score, as in Exps.1-3. Unlike results found in the other experiments, however, here we found a difference between critical and perceptual comparisons only in rating scores - $t(11) = -5.1916$, $p < 0.01$: perceptual comparisons (M=4.31, SD=1.03) were judged more similar than critical ones (M=2.65, SD=0.80); no significant difference in RTs was seen - $t(11) = 0.7237$, $p = 0.48$: judgments for critical (M=1336.96ms, SD=615.43ms) and perceptual (M=1196.31ms, SD=528.34ms) comparisons were given at a roughly equal speed.

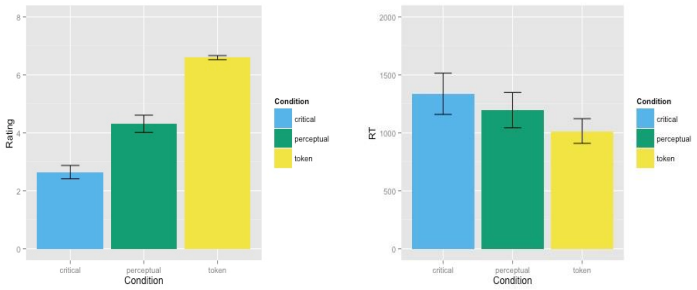

Figure 6: Mean ratings (left) and RTs (right) by condition in Experiment 4.

Discussion. A similarity judgment task with adults displayed sensitivity to the participation of a third entity in a variety of scene types. Sensitivity was revealed in either of two ways. For some types of stimuli (JIMMY/OPEN, STEAL/PICKUP), the relevant measure was RT, as we found with GIVE/HUG in Exp.1. For others, the relevant measure was rating score (BEAN/HIT). We will not speculate about what accounts for the different expressions of sensitivity across scene types. For now, what is important is that we can feel confident in using these materials to investigate infants' nonlinguistic representation of participanthood. With a clearer sense of this, we can ask what role participant representations play in verb learning, specifically with respect to PAM.

6 Preliminary results for infants

With the results from the adult experiments (Exp.1-4), we are confident that the materials we de-

veloped are appropriate for the purpose of assessing event perception. We therefore used them to probe event representation in prelinguistic infants. We now report some preliminary results for 10-month-olds.

First, we replicated Gordon (2003)'s finding that infants dishabituated to the disappearance of the teddy bear from giving type scenes, but not hugging type scenes. This suggests that infants do not view the hugging under a concept that has the role of the teddy as a participant role. The reason is, changes in a role that is psychologically foregrounded would presumably be noticed. Conclusions about the giving, where children do notice the disappearance of the teddy, are more delicate. There are many possible explanations for their noticing. The most plausible, we think, is that they view the scene under a concept that foregrounds the role of the object given. The two videos (hugging and giving) are objectively very similar; and yet they noticed the disappearance only in the giving video. A difference in the event representation is the best explanation of this difference.

Next, we deployed the same method with the jimmying scenes from Exp.2. We habituated infants to scenes in which a woman opens a box with a lever, from the left side. We then tested infants with a critical comparison (opening the box without a lever, but still from the left) and with a perceptual comparison (opening the box with a lever from the right). Infants dishabituated only to the former. We also habituated a different group of infants to scenes in which the opening is performed without a lever. Again, we tested these infants with a critical comparison (now opening the box with a lever) and with a perceptual comparison (opening from the other side). And again, they dishabituated only in the critical comparison.

These results suggest that infants, like adults, viewed these scenes under a concept that distinguishes the role of the lever as a participant. In their mental representation of these scenes, the lever is an explicit constituent, just like the eyes in our representation of a smiley human face. If this suggestion is correct, then the events of jimmying in our stimuli are viewed under a 3-participant concept. And yet, sentences we normally use to describe such scenes will often have only 2 arguments. PAM therefore predicts that children should have difficulty learning verbs like *jimmy*. We are currently testing this prediction.

7 Conclusion

PAM is a strong hypothesis about the way learners use sentence structure to inform their hypotheses about the concept that a given verb labels. On this view, event concepts have psychologically foregrounded participant roles and these participant roles are mapped one-to-one to syntactic arguments. Thus, the learner can use the number of syntactic arguments in a clause to (a) identify the number of participants in the event concept labeled by its verb and (b) to use that number *n* as a way of restricting candidate verb meanings: the clause, and hence the verb, applies only to events that are readily viewed under an *n*-participant concept. As discussed above, if it turns out that any event concept expressible as an *n*-participant concept can also be flexibly represented as an *n*-1 participant concept, then a learning mechanism based in PAM offers little guidance for learners. Thus, we take it to be an important goal to identify the participant structure of events independent of the verbs used to describe them.

In the current paper, we have begun an investigation into non-linguistic event representations so that the claims of PAM can be tested under a broader range of conditions than it has been tested under in the past. Specifically, we wish to identify scenes that are very likely to be viewed under an *n*-participant concept. We can then go on to examine whether such scenes are easily paired with simple transitive clauses. To the degree that they can, PAM is undermined. To the degree that they cannot, PAM is supported.

This paper identifies a range of cases that adults seem to readily represent as having three participants, but which can naturally be described using simple transitive clauses. We found that our scenes of jimmying, stealing and beaning are viewed under concepts that make the lever, victim and projectile participants. Interestingly, we sometimes found that this was revealed in explicit judgments and sometimes in the time it took to render this judgment.

These findings now set the stage for investigation of infant verb learners. We have already found that 10-month-olds seem to represent jimmyings like adults do, treating the lever as a psychologically privileged role. Investigations of thefts and beanings are underway. These results make us more confident about how children experience the world of the experiment, and so we

now find ourselves on stronger footing for testing PAM. To the degree that learners can pair a transitive clause containing a novel verb with a scene view under a 3-participant concept, this will raise questions about the utility of PAM as a learning bias. In turn, this will call for reconsideration of the biases that support inferences from argument number to verb meaning.

References

Casasola, Marienella and Leslie B. Cohen. 2000. Infants' association of linguistic labels with causal actions. *Developmental Psychology* 36(2): 155-168.

Fillmore, Charles. 1982. Frame Semantics. In *Linguistics in the morning calm: Selected papers from SICOL-1981*, ed. the Linguistic Society of Korea. Seoul: Hanshin Publishing Company.

Fisher, Cynthia. 1996. Structural Limits on Verb Mapping: The Role of Analogy in Children's Interpretation of Sentences. *Cognitive Psychology* 31: 41-81.

Goldberg, Adele. *Constructions*. Chicago: The University of Chicago Press.

Gordon, Peter. 2003. The origin of argument structure in infant event representations. In *Proceedings of the 28th Boston University Conference on Language Development*, ed. By A. Brugos, L. Micciulla, and C.E. Smith: 189-198.

Gleitman, Lila. 1990. The Structural Sources of Verb Meanings. *Language Acquisition* 1(1), 1990: 3-55.

Naigles, Letitia. 1990. Children use syntax to learn verb meanings. *Journal of Child Language* 17: 357.374.

Werker, Janet F., Leslie B. Cohen, Valerie L. Lloyd, Marianella Casasola, and Christine L. Stager. 1998. Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, *34*(6), 1289.

Williams, Alexander. 2015. *Arguments in Syntax and Semantics*. Cambridge: Cambridge University Press.

Department of Linguistics
Northwestern University
Evanston, IL 60208
wellwood@northwestern.edu

Department of Linguistics
University of Maryland
College Park, MD 20742–7505
angelahe@umd.edu
jlidz@umd.edu
alxndrw@umd.edu