Agent Preference in Children: The Role of Animacy and Event Coherence

Ceren Barış (ceren.baris@ozu.edu.tr)
Özyeğin University, Istanbul, Turkey

Ercenur Ünal (Ercenur.Unal@mpi.nl)
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Özyeğin University, Istanbul, Turkey

Abstract

Thematic roles in language (Agents, Patients) are considered to be hierarchically organized in terms of their salience, and this hierarchy is rooted in their counterparts as event participants in cognition. Here, we examine the relative salience of Agents over Patients in two-participant causative events in Turkish-speaking 3- to 5-year-old children. We also test if this asymmetry is modulated by the animacy of the Patient (human vs. inanimate object) and specific to the presence of a coherent event. In an eye-tracked change detection task, changes to Agents were detected more accurately (and after fewer fixations) than changes to inanimate Patients when there was a coherent event. This asymmetry disappeared when the Patient was animate (for accuracy) and when event coherence was disrupted (for both accuracy and fixations). These findings suggest an interplay of event roles and animacy in Agent preference.

Keywords: Event cognition; Thematic Hierarchy; Agent prominence; animacy; eye tracking

Introduction

A central part of human cognition is the ability to understand the events that unfold around us. When we view an event, we have to rapidly determine the relations that exist between the people, objects and entities involved in the event—in other words the *event participants*. For example, when we see a man knocking down some boxes, we need to conceptualize the man as the causer of the action (or *Agent*) and the boxes as the entities being affected by the action (or *Patient*). Research with adults has shown that viewers are able to rapidly extract event participants from briefly presented event visuals (Hafri et al., 2013; 2018). Further work has consistently shown that the ease with which different event participants are extracted varies (Isasi-Isasmendi et al., 2023; Ünal et al., 2024). However, less is known about how children extract event participants and whether they are also sensitive to the asymmetries between the salience of event participants. A further underexplored issue is how certain conceptual and perceptual factors characterize such asymmetries. Here, we investigate whether children show role prominence asymmetries for Agents over Patients in causative events and whether this asymmetry is modulated by the animacy of the Patient.

Event Participants in Cognition

According to an influential proposal, event participants closely correspond to the notion of thematic roles in language (Dowty, 1991; Jackendoff, 1990; see also Rissman & Majid, 2019). Thematic roles (e.g., Agent, Patient) capture the relationship between a verb and its constituents and relational information about event structure (i.e., 'who-did-what-to whom') (Fillmore, 1968). Furthermore, thematic roles are organized hierarchically: Agents are considered as the most prominent, followed by Patients, and then more peripheral roles (e.g., Goals, Sources, Instruments) (Baker, 1997; Jackendoff, 1990). Even though this formulation is based on linguistic behavior of thematic roles, there is growing evidence for the idea that thematic roles have counterparts in cognition and therefore role prominence asymmetries are reflected in conceptual prominence of event participants (Ünal et al., 2021a; 2024; see Jackendoff, 1990 for the initial formulation).

One line of support for this idea comes from studies investigating the rapid identification of event participants relevant for linguistic description of events (e.g., Griffin & Bock, 2000). In one study, participants viewed a wide range of two-participant events for 37 or 73 ms. After these brief viewings, they could reliably recognize the event, as well as who the Agent and the Patient were (Hafri et al., 2013). A follow-up study showed that Agent and Patient roles can be extracted spontaneously even when it is irrelevant to the task (Hafri et al., 2018). Further, adults categorize events based on who the Agent and the Patient is (Rissman & Lupyan, 2022).

Another class of studies revealed an asymmetry between the identification of event participants. In one study, participants saw briefly presented (100-300 ms) pictures of possession-transfer events (Dobel et al., 2007). Participants identified Agents with higher accuracy than the entities receiving the action, but this asymmetry was reduced when the coherence of the event was disrupted by mirroring the actors in the pictures. A later study used an eye-tracking paradigm to examine Agent prominence in Basque and Spanish speakers (Isasi-Isasmendi et al., 2023). Participants were briefly exposed to two-participant, Agent-Patient causative events and either described the event or answered a probe recognition question. Participants' first fixations were more likely to land on the Agent than the Patient. Furthermore, event descriptions were more specific and probe recognition was faster for Agents as opposed to Patients. These findings support the generalizability of Agent prominence over different classes of events and speakers of different languages.

Finally, a recent study investigated role prominence asymmetries in more complex caused motion events that involve additional roles (Ünal et al., 2024). In a linguistic

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

description task, participants mentioned (animate) Agents most frequently, followed by Patients, then Goals, and finally Instruments. This asymmetrical pattern was replicated in a non-linguistic visual search task measuring how fast people identified these event participants, showing evidence for parallels between how salient event roles are in language and cognition.

Developmental Evidence These correspondences between linguistic and conceptual organization of the internal structure of events raise questions about the developmental continuity of these correspondences. In fact, a number of studies inspired by the linguistic organization of events showed evidence for developmental origins of these correspondences in young infants (Göksun et al., 2008; Wagner & Lakusta, 2009). For instance, 7-month-old infants who habituated to two participant Agent-Patient events dishabituated when the roles of Agent and Patient were reversed (Leslie & Keeble, 1987). A recent study replicated these results using pupil dilation (Papeo et al., 2024). These findings suggest that some sensitivity to the internal structure of events and their participants is present by 7 months. Note that these studies indicate that infants have some understanding of Agent and Patient roles, not necessarily demonstrate a higher salience of Agents over Patients.

A recent study tested role prominence asymmetries among learners of English and Turkish (Ünal et al., 2021b). Children saw pictures of events depicting an animate Agent moving an object (Patient) to a Goal endpoint using an Instrument and were asked to indicate which event participant changed color. In both language groups, changes to both Patients and Goals were detected more accurately than changes to Instruments. Changes to Agents were also detected equally accurately in both language groups. Importantly, children did not have high levels of accuracy for Agent changes. This may not be completely unexpected given that for Agents only a part of the entity (i.e., clothes) changed color, whereas for the remaining participants the whole object changed color (see Ünal et al., 2024 for similar findings with adults). However, due to this difference in the nature of the color change and the difference in the animacy status, the salience of Agents was not compared to the remaining event participants. Even though these findings reveal some insights into the developmental continuity of role prominence asymmetries, both the nature of the color change and the difference in the animacy status of Agents vs. other roles introduce limitations to the study.

The Present Study

The work reviewed above leaves open several questions. One issue is the differences in the animacy status of event participants. In everyday events, Agents are usually animate, whereas Patients can be animate or inanimate. The visual system is attuned to this distinction (Konkle & Caramazza, 2013), prioritizing animate over inanimate entities (Thorpe et al., 2001). It is important to consider this factor given that previous evidence on the role of animacy on Agent prominence is mixed (Brocard et al., 2024; Isasi-Isasmendi et al., 2023).

A second issue concerns methodological factors. Recall that in previous work, the nature of the color changes in the change detection tasks was different for Agents (i.e., parts of objects) and Patients (i.e., whole objects). Given that this factor prevented previous work from making secure inferences about Agent prominence in children (and to some extent in adults), it is essential to use comparable color changes for Agents and Patients.

Finally, it is important to ensure that the asymmetries in the relative salience of event participants are indeed driven by the conceptual role they play in the event and not due to the visual features (e.g., shape, color, complexity, size) of the people or objects that fill these roles. A promising method to explore this issue involves disrupting event coherence by mirroring the people or objects in the events and using these back-to-back (non-)event versions as controls (cf. Dobel et al., 2007). Previous work shows that these back-to-back (non-)event versions indeed have lower dyadic coherence compared to their face-to-face counterparts (Goupil et al., 2023; Papeo et al., 2017), highlighting the potential of this method to rule out visual features as confounds.

In the present study, we address the open questions on Agent preference. We ask whether preschool-aged children show prominence asymmetries for Agents over Patients in causative scenes and whether this is modulated by the animacy of the Patients. We also ask if any difference we find in saliency is indeed due to the conceptual roles of event participants and cannot be attributed to other (low-level) visual features.

We used an eye-tracked change detection paradigm. Our stimuli were composed of two-participant event depictions with either an animate or an inanimate Patient. New versions of these depictions were created by applying a color change to either the Agent or the Patient. The stimuli set also included 'non-event' versions for which each event participant's vertically symmetrical image was used so they were back-to-back (see Figure 1). We measured both the accuracy of change detection and recorded children's eye movements as an index of the time course of the identification of the changing participant. We aimed to determine whether change detection accuracy and fixations to the changing participant varied based on event role, animacy of the Patient, and event coherence.

We hypothesized that children would demonstrate a stronger Agent-Patient asymmetry in face-to-face events that include inanimate Patients, but this asymmetry should diminish or disappear in face-to-face events that include animate Patients. Thus, changes to Agents should be detected more accurately than changes to (inanimate) Patients. Further, fixations to the changing participant should differ across Agents and (inanimate) Patients. Finally, we expected these Agent-Patient asymmetries to completely disappear for back-to-back (non-)events regardless of the animacy of the Patient.

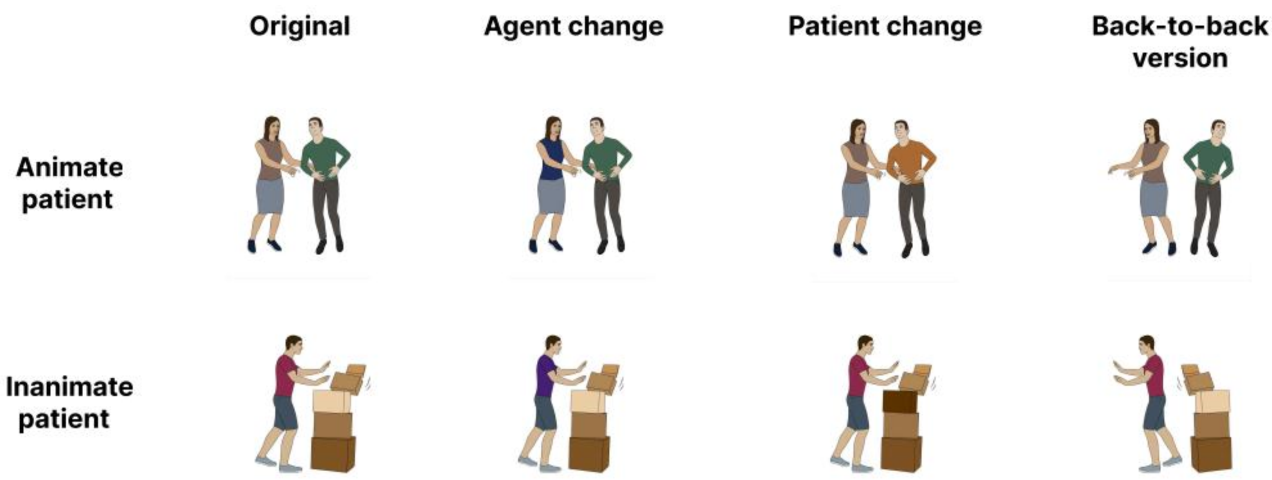

Figure 1: Examples of stimuli.

Method

Participants

Data were collected from preschoolers (n = 33, 25 females, M_{age} = 4.67 years, SD_{age} = 0.56, range = 3.49 – 5.42). All were native speakers of Turkish. Participants were recruited from various preschools in Istanbul, Turkey. Fourteen additional preschoolers were tested but excluded from the analysis due to not following the instructions (n = 2), not completing the experiment (n = 1), having less than 50% accuracy (n = 4), or being older than 5.5 years (n = 7). The sample size was determined based on power analysis.

Materials

Stimuli were created using Procreate and edited using Adobe Photoshop. Target stimuli consisted of 12 illustrations that depicted midpoints of causative events in which an Agent acted on a Patient on a white background (see Figure 1). The Agent was always human. The Patient was human in half of the target events and an inanimate object in the other half. When selecting the events and creating the illustrations with inanimate Patients, we ensured that the objects were similar in size to humans to eliminate the possible effect of size on salience. Female-male status of the Agents and Patients was equally distributed and counterbalanced across items. Agents and Patients faced each other in all the illustrations. We created new versions of the images, in which part of the Agent's clothing color was different. The same was done for the Patients (for inanimate Patients, parts of the objects changed color).

We used back-to-back (non-)event versions to disrupt the event coherence while keeping the visual features such as body posture, size, and relative complexity the same. Since the back-to-back (non-)event versions do not depict coherent events, the entities in them are not event participants. Nevertheless, for ease of reference, we will refer to them as Agents or Patients. A pilot study confirmed that face-to-face events were rated significantly higher as depicting coherent events than back-to-back (non-)events ($t(30.765) = -8.493$, $p < .001$).

Procedure

Participants were tested individually in a quiet room in their preschool. Participants were seated approximately 60 cm away from a DELL Precision M4800 laptop with the SMI RED 250 eye-tracker (SensoMotoric Instruments) mounted underneath the screen. The stimuli were presented via NBS Presentation software (Version 23.1, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

Each version of each event (change in Agent or Patient and face-to-face or back-to-back) was seen by four different participants, but each participant saw each event only once. Further, each participant saw 3 instances of Agent and 3 instances of Patient change in either category. The side of the color change was counterbalanced within events. Due to an error in counterbalancing lists, face-to-face and back-to-back (non-)events were distributed unevenly across participants, but the other variable levels were distributed evenly. Each participant saw a total of 12 trials and 6 filler (noncausative) trials. The trials were presented in a single randomized order.

The change blindness procedure was adapted from Rensink and colleagues (1997) (for studies with children see Ünal et al., 2021b; Shore et al., 2006). Each trial began with a fixation cross that was presented in the middle of the screen for 250 ms. Then, the original event appeared on the screen for 240 ms, followed by a grey screen mask that appeared for 80 ms. After the mask, the color-changed version of the event appeared for 240 ms, again followed by a mask for 80 ms. This loop was repeated 7 times based on a pilot study. Thus, each trial lasted 4.48 seconds. At the end of each trial, participants were asked on which side the color change was. They were instructed to raise their left or right hand based on the color change side. This procedure was also tested in a pilot study.

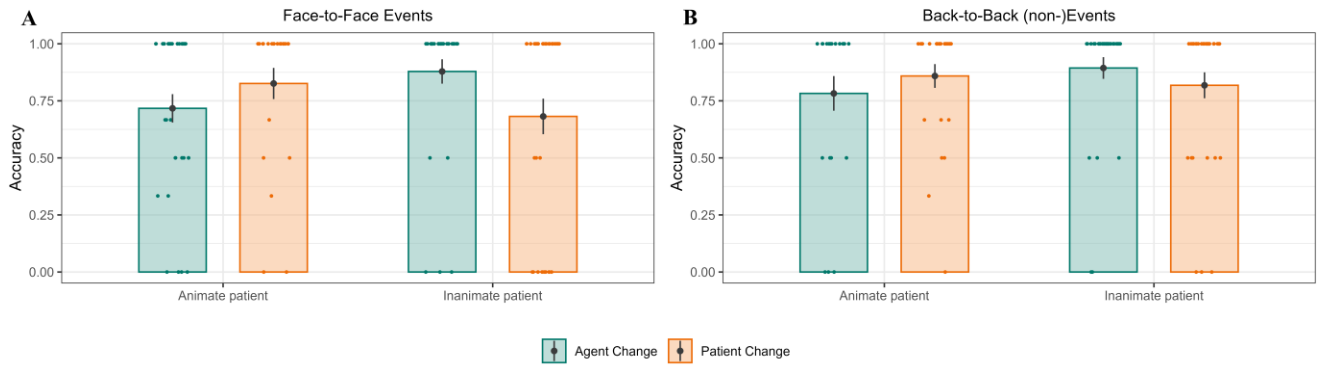

Figure 2: The height of the bars depicts average accuracy across changing role and animacy of the Patient in (A) face-to-face events and (B) back-to-back (non-)events. Error bars indicate standard error.

The experiment began with two practice trials, which depicted two cartoon animals standing opposite each other, one of which changed color. Then, the participants completed a 5-point calibration and validation procedure. Participants were given an opportunity to ask questions before they began the main task with 18 trials. The entire task took 10 minutes.

Results

We first examined whether accuracy varied based on the changing role and animacy of the Patient. We then examined the variation in participants' target fixations depending on the changing role and animacy of the Patient.

Accuracy Data

Four participants had less than 50% accuracy throughout the entire task and were excluded from further analyses. All exclusions were done prior to data analysis. Accuracy data were analyzed using generalized binomial linear mixed effects models. Models were fit using the *glmer* function in the *lme4* package (version 1.1.33; Bates et al., 2015) in R (version 4.2.3; R Core Team, 2023).

Face-to-Face Events We first analyzed the face-to-face trials to determine whether Agent preference emerged when there was a coherent event. The model tested the fixed effects of changing role and animacy of the Patient on the binary dependent variable of accuracy (1=correct, 0=incorrect) at the trial level (see Figure 2A). Fixed effects were tested with sum-to-zero contrasts (-0.5, 0.5) (Schad et al., 2020). The model included random intercepts for Subjects (Baayen et al., 2008). The inclusion of random intercepts for Items produced a singular fit error, hence, this term was excluded from the model.

The model revealed a significant interaction between changing role and animacy of the Patient. In the inanimate Patient condition, participants were more accurate when detecting changes to Agents than Patients ($\beta = 1.2743$, $SE = 0.6158$, $z = 2.069$, $p = .04$). However, in the animate Patient condition they were equally accurate in detecting changes to Agents and Patients ($\beta = -0.8977$, $SE = 0.5445$, $z = -1.649$, $p = .10$).

Back-to-Back (non-)Events We then examined the back-to-back trials to make sure any difference we found in face-to-face trials was due to event roles and not visual features. The model tested the fixed effects of changing role and animacy of the Patient on the binary dependent variable of accuracy at the trial level (see Figure 2B). Fixed effects were tested with sum-to-zero contrasts (-0.5, 0.5) (Schad et al., 2020). The model included random intercepts for both Subjects and Items (Baayen et al., 2008).

No fixed effects or interactions were statistically significant. Participants were equally accurate at detecting changes to Agents and Patients in back-to-back (non-)events both with an animate or an inanimate Patient.

Eye Gaze Data

Participants' fixations to Agent and Patient areas of interest (AoI) were computed using the SMI BeGaze software. We analyzed eye gaze data only for accurate trials (81% of the data). No participant had more than 45% trackloss across all trials. 15% of the accurate trials had more than 50% trackloss and were excluded from further analysis. All exclusions were done prior to data analysis.

We then computed the proportion of fixations to changing event participant (target) out of all fixations to the screen. From this, we aggregated the proportion of target fixations in 1000 ms time windows for the first 2000 ms of the trial. We reasoned that eye movements during the earlier time windows would more precisely reflect the visual identification of target event roles since event participants are detected rapidly (Hafri et al., 2013; 2018) and children fixate on a changing target in 1000 ms (e.g., Hirose & Hancock, 2007). Eye gaze data were analyzed using linear mixed effects models. Models were fit using the *lmer* function in the *lme4* package (version 1.1.33; Bates et al., 2015) in R (version 4.2.3; R Core Team, 2023). As in the analysis of the

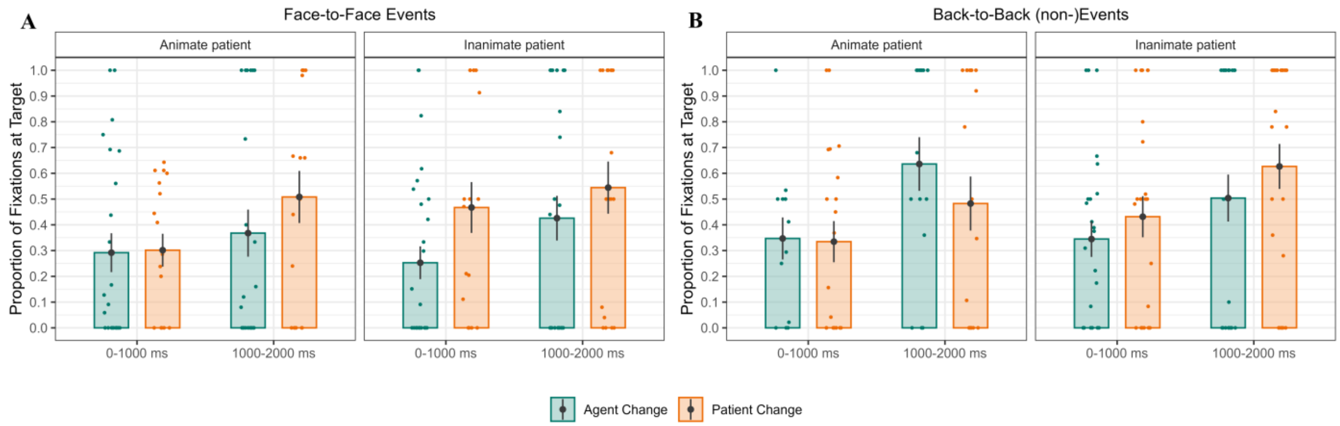

Figure 3: The height of the bars depicts average proportion of fixations at target across changing role in each time window across animacy of the Patient for (A) face-to-face events and (B) back-to-back (non-)events. Error bars indicate standard error of participant means.

accuracy data, data from face-to-face events and back-to-back (non-)events were analyzed separately.

Face-to-Face Events The model tested the fixed effects of changing role, animacy of the Patient and time window (0-1000 ms, 1000-2000 ms) on the dependent variable of proportion of target fixations at the trial level (see Figure 3A). The fixed effects were tested with sum-to-zero contrasts (-0.5, 0.5) (Schad et al., 2020). The model included random intercepts for Subjects (Baayen et al., 2008). The inclusion of random intercepts for Items produced a singular fit error, hence, this term was excluded from the model.

The model revealed only a significant main effect of changing role on the proportion of target fixations ($\beta = 0.117$, $SE = 0.053$, $t = 2.181$, $p = .03$). Participants fixated more on the target when detecting changes to Patients compared to when detecting changes to Agents, regardless of Patient animacy.

Back-to-Back (non-)Events The model tested the fixed effect of changing role, animacy of the Patient and time window on the dependent variable of proportion of target fixations at the trial level (see Figure 3B). The fixed effects were tested with sum-to-zero contrasts as the face-to-face model (Schad et al., 2020). The model included random intercepts for Subjects and Items (Baayen et al., 2008).

There were no main effects of changing role or Patient animacy, nor any interactions. Participants fixated on the target equally when they were detecting changes to Agents and Patients in back-to-back (non-)events, both with an animate or an inanimate Patient.

Discussion

In the present study, our main goal was to assess whether there was an asymmetry between the relative salience of Agents and Patients in causative events and whether this asymmetry was modulated by the animacy of the Patient. We also explored whether this asymmetry was indeed due to the presence of a coherent event and could not be attributed to variations in the low-level visual features of the people or objects filling these event roles. To do so, we used the same type of color change for Agents and Patients and manipulated the presence of a coherent event while controlling for visual complexity.

In line with our expectations, when there was a coherent event, children were more accurate in detecting changes to Agents as opposed to inanimate Patients. However, the asymmetry between Agents and Patients disappeared when the Patient was animate and when there was no coherent event. These findings are consistent with previous work with adults showing a higher relative salience of Agents over Patients (Dobel et al., 2007; Isasi-Isasmendi et al., 2023; Ünal et al., 2024). Importantly, current results point to an effect of animacy in the relative salience of Agents and Patients. It is possible that an event with one animate and one inanimate entity provides an easier frame to assign Agent/Patient roles through animacy signaling agency.

Turning to the eye gaze data, when there was a coherent event, children had more target fixations and hence took longer to accurately detect changes to Patients as opposed to changes to Agents. This suggests that children needed to glean more information from Patients or to process them for longer to be able to accurately detect the changes to them, presumably because they were less salient. This is in line with our prediction that target fixations would differ when detecting changes to Agents versus Patients. This interpretation of the eye gaze data is also consistent with previous developmental studies using eye gaze based measures that have interpreted increased looking times as an indicator of longer cognitive processing (e.g., Birulés et al., 2024; Candan et al., 2012). Together with the accuracy results, these findings suggest that the difference in salience of Agents and Patients is modulated by the animacy status of the Patient but not driven by low-level visual features.

One might think that the direction of the difference between Agents and inanimate Patients in the eye gaze data is at odds with the findings of previous work with adults relying on eye gaze measures as an index of relative salience of event participants (e.g., Griffin & Bock, 2000; Ünal et al., 2024). In that work, more salient entities were allocated more visual attention. Nevertheless, these studies have used visual search paradigms, instructing participants to find an event component (e.g., the object affected by the action) and keep looking at it until a button press response. Because we did not instruct our participants to keep looking at the changing component, they were free to look at other parts of the visual stimuli once they detected which person or object was changing color. This methodological difference might explain the two sets of seemingly conflicting findings. On the other hand, our findings are consistent with previous work using change detection tasks, showing that accurate detection of changes to less salient event participants was associated with longer reaction times (Ünal et al., 2024; see also Rensink et al., 1997). This suggests that combining eye tracking with change detection paradigms might be a promising approach to gain insights into the temporal dynamics of change detection in preschool-aged children. Nevertheless, data from adults from similar paradigms is needed both to complete the developmental picture and to establish the validity of eye gaze based measures of change detection. We are currently addressing this in ongoing work.

A notable difference between the accuracy and eye gaze data concerns the effect of the animacy of the Patient. As previously discussed, the accuracy results indicated higher salience of Agents compared to inanimate Patients but not compared to animate Patients. In contrast, the eye gaze data showed a consistent Agent prominence over Patients regardless of animacy. This discrepancy suggests that animacy may not be a reliable cue for agency in the early stages of children's event processing. Instead, the link between animacy and agency may emerge through experience during development, being applicable only at later stages of processing in preschool-aged children. Work with older children or adults can help evaluate this explanation more directly.

Viewed from a broader perspective, our findings are consistent with the view that linguistic organization of events builds on the structure offered by cognition to capture the internal representation of events (Jackendoff, 1990, cf. Rissman & Majid, 2019; Ünal et al., 2021a). From a developmental perspective, our findings connect to a larger literature on how children learn verbs labelling different types of events (e.g., Gleitman, 1990; Pinker, 1989; Tomasello & Merriman, 2014). This work has shown that children are able to keep track of the connections between verb meanings and event structure, the number of event participants, and the relations that exist between them (Arunachalam & Waxman, 2010). Further, studies show that 2-year-olds can resolve ambiguous pronouns by favoring subjects over objects as referents (Song & Fisher, 2005; 2007). This linguistic subject preference may reflect an underlying cognitive bias for Agents, as observed in the present study. Nevertheless, a large portion of this work includes children who are younger than the present age group of preschoolers. One challenge for future work is bridging the gap between the work with infants and preschoolers to be better able to characterize how verbs are mapped onto event referents in language acquisition and how this might be shaped by cognitive biases such as the salience of Agents.

In conclusion, the current study highlights the interplay of event roles and animacy in Agent prominence. By examining Agent-Patient asymmetries in a non-Indo-European language that allows omission of subjects and with preschool-aged children, our findings contribute to the cross-linguistic and developmental continuity of Agent prominence.

Acknowledgments

This research was supported by grant 121K259 awarded to E.Ü. and 2210-A graduate fellowship awarded to C.B. by the Scientific and Technological Research Council of Turkey.

References

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442-446. https://doi.org/10.1016/j.cognition.2009.10.015

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baker, M.C. (1997). Thematic roles and syntactic structure. In L. Haegeman (Ed.) *Elements of grammar: Handbook of generative syntax* (pp. 73–137). https://doi.org/10.1007/978-94-011-5420-8

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

Birulés, J., Bosch, L., Lewkowicz, D. J., Pons, F. (2024). Time course of attention to a talker's mouth in monolingual and close-language bilingual children. *Developmental Psychology*, *60*(1), 135-143. https://doi.org/10.1037/dev0001659

Brocard, S., Wilson, V. A., Berton, C., Zuberbühler, K., & Bickel, B. (2024). A universal preference for animate agents in hominids. *iScience*, *27*(6), 109996. https://doi.org/10.1016/j.isci.2024.109996

Candan, A., Küntay, A. C., Yeh, Y., Cheung, H., Wagner, L., & Naigles, L. R. (2012). Language and age effects in children's processing of word order. *Cognitive Development*, *27*(3), 205-221. https://doi.org/10.1016/j.cogdev.2011.12.001

Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*(2), 129–143. https://doi.org/10.1016/j.actpsy.2006.07.004

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619. https://doi.org/10.2307/415037

Fillmore, C. J. (1968). The case for Case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). Holt, Rinehart and Winston.

Gleitman, L. R. (1990). The structural sources of verb learning. *Language Acquisition*, *1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Goupil, N., Hochmann, J. R., & Papeo, L. (2023). Intermodulation responses show integration of interacting bodies in a new whole. *Cortex*, *165*, 129–140. https://doi.org/10.1016/j.cortex.2023.04.013

Göksun, T., Küntay, A. C., & Naigles, L. R. (2008). Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language*, *35*(2), 291–323. doi:10.1017/S0305000907008471

Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11* (4), 274–279. https://doi.org/10.1111/1467-9280.00255

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, *142*(3), 880–905. https://doi.org/10.1037/a0030045

Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher level visual processing. *Cognition*, *175*, 36–52. https://doi.org/10.1016/j.cognition.2018.02.011

Hirose, Y., & Hancock, P. J. B. (2007). Equally attending but still not seeing: An eye-tracking study of change detection in own- and other-race faces. *Visual Cognition*, *15*(6), 647–660. https://doi.org/10.1080/13506280601069578

Isasi-Isasmendi, A., Andrews, C., Flecken, M., Laka, I., Daum, M. M., Meyer, M., Bickel, B., & Sauppe, S. (2023). The agent preference in visual event apprehension. *Open Mind*, *7*, 240–282. https://doi.org/10.1162/opmi_a_00083

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *Journal of Neuroscience*, *33*(25), 10235-10242. https://doi.org/10.1523/JNEUROSCI.0983-13.2013

Leslie, A. M., and 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

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. P*sychological Science*, *28*(3), 369-379. https://doi.org/10.1177/0956797616685769

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. https://doi.org/10.1016/j.cub.2024.07.081

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. MIT Press. https://doi.org/10.7551/mitpress/4158.001.0001

R Core Team (2023). *R: A language and environment for statistical computing*. https://www.R-project.org/.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368–373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Rissman, L., & Lupyan, G. (2022). A dissociation between conceptual prominence and explicit category learning: Evidence from agent and patient event roles. *Journal of Experimental Psychology: General*, *151*(7), 1707–1732. https://doi.org/10.1037/xge0001146

Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, *26*, 1850–1869. https://doi.org/10.3758/s13423-019-01634-5

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10.1016/j.jml.2019.104038

Shore, D. I., Burack, J. A., Miller, D., Joseph, S., & Enns, J. T. (2006). The development of change detection. *Developmental Science*, *9*(5), 490-497. https://doi.org/10.1111/j.1467-7687.2006.00516.x

Song, H., & Fisher, C. (2005). Who's ‘‘she’’? Discourse structure influences preschoolers' pronoun interpretation. *Journal of Memory and Language*, *52*(1), 29–57. doi:10.1016/j.jml.2004.06.012

Song, H., & Fisher, C. (2007). Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua*, *117*(11), 1959-1987. https://doi.org/10.1016/j.lingua.2006.11.011

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience*, *14*(5), 869-876. https://doi.org/10.1046/j.0953-816x.2001.01717.x

Tomasello, M., & Merriman, W.E. (Eds.). (2014). *Beyond names for things: Young children's acquisition of verbs*. Psychology Press. https://doi.org/10.4324/9781315806860

Ünal, E., Ji, Y. & Papafragou, A. (2021a). From event representation to linguistic meaning. *Topics in Cognitive Science*, *13*, 224-242. https://doi.org/10.1111/tops.12475

Ünal, E., Richards, C., Trueswell, J. C., & Papafragou, A. (2021b). Representing agents, patients, goals and instruments in causative events: A cross-linguistic investigation of early language and cognition. *Developmental Science*, *24*(6), e13116. https://doi.org/10.1111/desc.13116

Ünal, E., Wilson, F., Trueswell, J., & Papafragou, A. (2024). Asymmetries in encoding event roles: Evidence from language and cognition. *Cognition*, *250*, 105868. https://doi.org/10.1016/j.cognition.2024.105868

Wagner, L., & Lakusta, L. (2009). Using language to navigate the infant mind. *Perspectives on Psychological Science*, *4*, 177–184. https://doi.org/10.1111/j.1745-6924.2009.01117.x