Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative

Neil Cohn [a,*], Martin Paczynski [b]

[a] Center for Research in Language, University of California, San Diego, La Jolla, CA, USA
[b] Department of Psychology, University of Miami, Coral Gables, FL, USA

ARTICLE INFO

ABSTRACT

Agents consistently appear prior to Patients in sentences, manual signs, and drawings, and Agents are responded to faster when presented in visual depictions of events. We hypothesized that this "Agent advantage" reflects Agents' role in event structure. We investigated this question by manipulating the depictions of Agents and Patients in preparatory actions in wordless visual narratives. We found that Agents elicited a greater degree of predictions regarding upcoming events than Patients, that Agents are viewed longer than Patients, independent of serial order, and that visual depictions of actions are processed more quickly following the presentation of an Agent vs. a Patient. Taken together these findings support the notion that Agents initiate the building of event representation. We suggest that Agent First orders facilitate the interpretation of events as they unfold and that the saliency of Agents within visual representations of events is driven by anticipation of upcoming events.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

When describing an event, people tend to mention the actor prior to the undergoer, or more formally, the Agent before the Patient. This ordering preference has been found across diverse languages (Kemmerer, 2012), suggesting that the order might be driven by a cognitive universal, a supposition

---

* Corresponding author. Address: Center for Research in Language, University of California, San Diego, 9500 Gilman Dr. Dept. 0526, La Jolla, CA 92093-0526, USA. Fax: +1 858 822 5097.
E-mail address: neilcohn@visuallanguagelab.com (N. Cohn).

supported by findings that similar order preferences appear in spontaneously generated gestures (Goldin-Meadow, So, Ôzyûrek, & Mylander, 2008) and when ordering drawings of events (Gershoff-Stowe & Goldin-Meadow, 2002). This "Agent advantage" also appears when people view visual depictions of actions, with people preferentially orienting to Agents over Patients (Germeys & d'Ydewalle, 2007) and identifying Agents faster than Patients. In the current series of experiments we sought to explore the reason for this apparent Agent advantage using serial visual narratives familiar to most people: comic strips. Below we first review previous literature concerning the Agent advantage. We then discuss why visual narratives offer a unique way of exploring the psychological underpinnings of the phenomenon and how this methodology can overcome the limitations of previous investigations.

1.1. Serial ordering of semantic roles

The canonical serial order of semantic roles within grammatical constructions places Agents prior to Patients, as expressed in SVO or SOV constructions across languages. Linguistic theories have long placed a preference on Agents over Patients (Clark, 1965; Jackendoff, 1972, 1990; Silverstein, 1976; Zubin, 1979). Agents are given the highest ranking in most hierarchies of semantic roles (e.g., Aissen, 1999; Jackendoff, 1972; Silverstein, 1976), and some work has proposed a universal constraint that selects Agents as the default grammatical subject of sentences (Keenan, 1976). Indeed, Agent–Patient orders are far more common as the canonical pattern in the world's languages than the reverse (Dryer, 2011; Greenberg, 1966; Hawkins, 1983).

Psycholinguistic research spanning several decades has provided evidence in support of a canonical serial ordering of thematic roles playing an important part in comprehension. The placement of Patients first, rather than Agents, has been shown to render sequences harder to understand (e.g., Ferreira, 2003; Hoeks, Vonk, & Schriefers, 2002; Miller & McKean, 1964; Tannenbaum & Williams, 1968). For example, Ferreira (2003) found that participants were slower and less accurate in correctly identifying Agents and Patients in passive sentences, compared to active ones. Ferreira demonstrated this effect even for infrequent active constructions, suggesting that the effect was driven by canonical semantic role ordering rather than an artifact of syntactic frequency. These findings were interpreted as supporting the use of a so-called noun–verb–noun (NVN) heuristic in mapping arguments to thematic roles, with the first encountered noun being interpreted as the Agent and the second as the Patient. Because the heuristic is independent of specific syntactic constructions, the proposal easily adapts to non-linguistic domains.

Research of manual communication further supports a canonical serial order. An Agent–Patient-Act pattern consistently appears in the "homesign" systems created by deaf children who are not taught a sign language (e.g., Goldin-Meadow, 2003; Goldin-Meadow & Feldman, 1977). Since the children's signs are often more extensive and complex than their parents', it is thought that it is the children, rather than caregivers, who are driving the creation of these homesigns (Goldin-Meadow & Mylander, 1983). This Agent–Patient-Act order appears in homesigns across the world, regardless of a culture's spoken languages (Goldin-Meadow et al., 2008). Interestingly, a similar pattern emerges when speaking individuals communicate through gesture alone (Gershoff-Stowe & Goldin-Meadow, 2002), a finding that appears to be independent of the language spoken by the participants (Goldin-Meadow et al., 2008). Taken together, these findings lend further credence to the idea that the canonical serial ordering of arguments may be a more general aspect of human "event cognition."

Canonical serial order effects have also been observed in the visual-graphic domain. Goldin-Meadow and colleagues asked speaking adults to view animated scenes of characters and objects doing simple actions. When the participants reconstructed these events using only images, they consistently produced an Agent–Patient-Act pattern (Gershoff-Stowe & Goldin-Meadow, 2002), with this pattern, again, being independent of the different syntactic constraints of a speaker's native language (Goldin-Meadow et al., 2008). Because of these findings, Goldin-Meadow has hypothesized that this Agent–Patient-Act order may be the natural way of comprehending—or at least expressing—events (Gershoff-Stowe & Goldin-Meadow, 2002; Goldin-Meadow et al., 2008).

Effects of canonical serial order also appear in correlations between semantic roles and spatial orientations (for review see Chatterjee, 2001, 2010). When people draw stick figures depicting various events, Agents are consistently drawn to the left of Patients (Chatterjee, Maher, & Heilman, 1995;

Chatterjee, Southwood, & Basilico, 1999), and similar orientations maintain when participants verbalize their visual imagery of events (Geminiani, Bisiach, Berti, & Rusconi, 1995). Also, faster reaction times have been found to matching sentences with pictures if the images show an Agent on the left as opposed to the right (Chatterjee et al., 1999). While Agents prototypically correspond to a leftward spatial position and Patients to the right, such correspondences appears to be dependent on the directionality of participants' writing systems (Dobel, Diesendruck, & Bölte, 2007; Maass & Russo, 2003). Thus, the prevalence of Agents on the left indicates a preference for Agents as *prior to* Patients for readers of left-to-right writing systems (and the reverse for users of opposite systems).

1.2. Beyond canonical serial order

Nevertheless, serial order does not seem to account for all instances of the Agent advantage. For example Segalowitz (1982; Experiment 1) presented participants with line drawings of two fish where one was biting the other. Pictures depicted either the left-most fish biting the right, the right-most fish biting the left, or neither fish biting the other. Measured with a button press, participants were able to identify which fish was the Agent (in one block) faster than which was the Patient (in the other block). Importantly, this Agent advantage was found to be independent of Agent position, suggesting that Agent processing was distinct from sequential semantic role processing (Experiment 3).

The Agent advantage has also been reported in studies with film. Children attend more to Agents than Patients while viewing an event, though attention is evenly distributed prior the event's start (Robertson & Suci, 1980). Studies with adults also have shown that attention becomes focused on the character with greater agency, no matter which side of the screen they are placed (Germeys & d'Ydewalle, 2007). This effect appears to be quite general, being observed not only when the entities observed are classic "Agents," such as humans. Agents are responded to faster than Patients in research using a classic paradigm where geometric shapes appear to have causal interactions with each other (Verfaillie & Daems, 1996), such as when circles ricochet off each other (e.g., Michotte, 1946). Finally, characters are also more often described as Agents than Patients by viewers of narrative films, especially if they play a more prominent role in the narrative than secondary characters (Magliano, Taylor, & Kim, 2005). Such findings suggest that the Agent advantage has cognitive foundations that may be independent of serial ordering and instead point to Agents having a central role in how we view events.

1.3. Agents as event builders

As discussed, the Agent initiates an action. In explaining the predominance of subject before object word ordering cross-culturally (89%), Kemmerer (2012) proposed a *subject salience principle* where iconicity drives this effect: because the prototypical Agent is the head of the causal chain responsible for inducing change in the Patient, the Agent is placed at the head of the construction. Put another way, because the Agent drives the event to take place to begin with, the Agent becomes the point of attention or anchor for the upcoming information. Thus, we propose that the canonical serial order of semantic roles within language, manuals signs, and other forms of expression may have the same processing basis as the faster identification of and orienting to Agents over Patients observed in film: the Agent serves to initiate the construction of an event representation followed by a subsequent linking of the actual event with Patients.

How can we provide evidence for this proposal? Studies of language make the question difficult to answer for several reasons. As noted above, the canonical Agent-before-Patient semantic role ordering is, necessarily, more frequent than non-canonical semantic role orderings. This makes it difficult to disentangle the effects of familiarity for semantic role ordering, especially if, as proposed by Ferreira (2003), the language parser utilizes a heuristic to interpret transitive constructions. Additionally, semantic roles in language can be difficult to determine prior to the presentation of the verb itself, the point at which the roles are formally syntactically assigned. This makes it difficult to separate the effects of semantic roles from effects of reanalysis, as might be the case if an argument initially presumed to be an Agent instead becomes assigned the Patient role, or vice versa.

Film likewise faces several confounds. First, in previous investigations utilizing film, the Agent was the first object to move, thus making it difficult to distinguish any Agent specific saliency from effects driven by change detection per se. Second, filmed actions usually conflate the Agent and the action being performed, thereby making it difficult to determine whether the Agency or the action drives any observed effects.

1.4. Visual narratives

Visual narratives consisting of sequential images, like those in comics, offer an alternative means of investigating linguistic questions while avoiding some inherent confounds. As laid out in Cohn (2013; Cohn, Paczynski, Jackendoff, Holcomb, & Kuperberg, 2012), sequential visual narratives share, functionally, many structural features analogous to syntax in language, despite operating at the discourse level. For example, visual narrative grammar uses recrusive principles that allow complete narratives to be coherently embedded within longer narratives, much as independent clauses can be embedded within a sentence. As argued in Cohn (2010, 2013), the comprehension of visual narratives requires hierarchical structuring of individual units into an overarching structure, similar to the hierarchical organization of words within a sentence. Importantly, within the theory of visual narrative grammar put forth by Cohn (2013), individual panels which make up a visual narrative combine, predictibly and systematically, to build event representations.

This framework of visual narrative grammar is particularly well-suited for addressing the question of the role that Agents and Patients play in people's understanding of events. In contrast to language, sequential presentation of semantic roles within visual narratives does not suffer from the same familiarity effects of serial order, and in contrast to film, Agents, like Patients, are necessarily static, thus eliminating the confound of saliency due to motion or change detection. More importantly, within visual narratives the Agent can be effectively separated from the depiction of the actual action, as can be seen in Fig. 1. Here, the panel depicting the *preparatory action* allows the comprehender to make expectations about characters' semantic roles prior to the actualization of those roles in the subsequent panel. What's more, each character depicting a semantic role can be individuated in their own panel (Cohn, 2007). Thus, visual narratives offer ways to avoid confounds seen in previous investigations of the role of Agents and Patients during event comprehension.
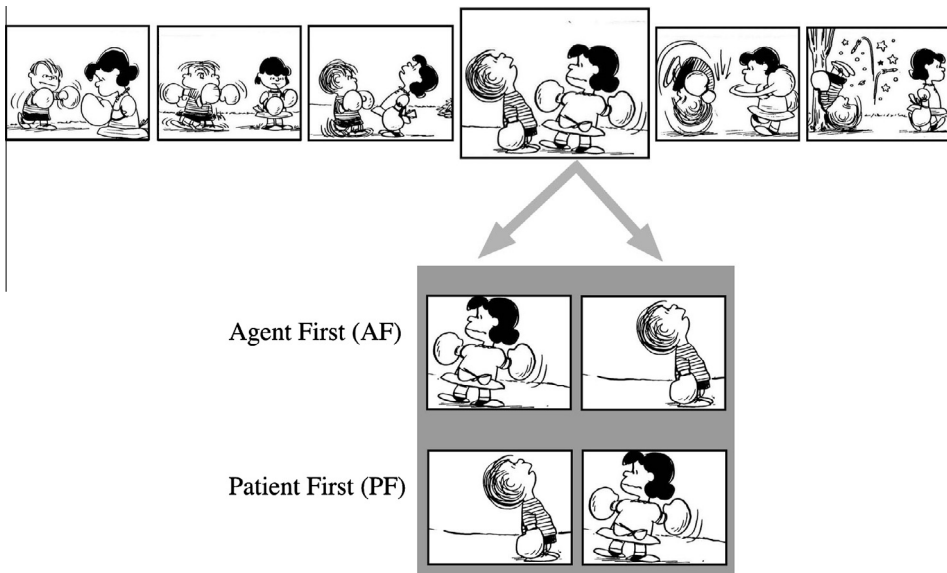

Fig. 1. Visual narrative with a panel depicting a preparatory action divided into individual panels, which keep each semantic role distinct. These independent panels then can feature alternating orders with either Agent or Patient presented first.
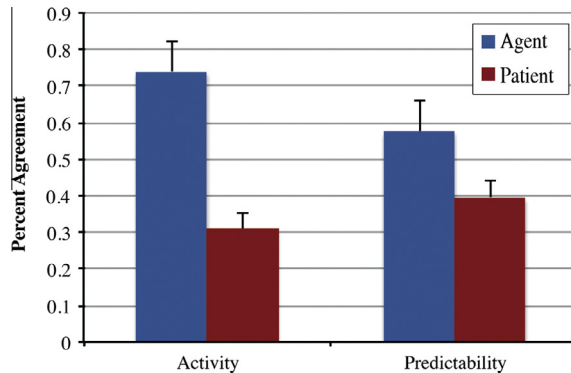
Fig. 2. Mean ratings for Activity (Experiment 1) and Predictability (Experiment 2) of Agent and Patient panels.

1.5. The present study

In this paper we present five experiments that investigate the Agent advantage previously reported in studies of language, gesture, film, and static images. In Experiment 1, we seek to validate our proposal that Agent and Patient roles can indeed clearly be differentiated prior to the actual execution of the action. In Experiment 2, we examine whether participants establish the foundation of the upcoming event based on Agents', but not Patients', involvement in preparatory actions. In Experiment 3, we explore whether altering the serial order of entities where the semantic roles are not yet established can influence the online viewing of panels in visual narratives. These first three experiments validate our experimental materials and test order-based effects in preparation for our final two core experiments. In Experiment 4, we conducted a "self-paced viewing study" to determine the impact of serial ordering on semantic role processing. In Experiment 5, we used self-paced viewing of sequences where only an Agent or a Patient preceded a depiction of the actual action to see how each semantic role impacts the processing of the completed event.

2. Experiment 1: Validation of semantic roles

We have proposed that the semantic role of an entity, in particular the Agent role, within visual narratives, such as comic strips, can be established *prior* to the action being executed. Previous experiments have shown that participants can reliably extract semantic roles from visually depicted events even in rapidly presented visual stimuli, and that such roles were identified based on particular semantic features (such as outstretched arms) (Hafri, Papafragou, & Trueswell, 2013). However, research has yet to confirm that people can identify semantic roles when presented in states *prior* to the completed actions. These preparatory actions undertaken by characters in the panels should provide cues that suggest they are either Agents or Patients prior to their completed actions. For example, in Fig. 1, Lucy reaching her arm back in preparation to punch should identify her as an Agent. Meanwhile, Linus's posture does not suggest that he is about to punch (despite also wearing boxing gloves), and indeed implies that he is distracted, which should identify him as a Patient. Participants should thus be able to make use of such cues to identify "expected" semantic roles of entities prior to their incorporation in the completion of the punching event.

In Experiment 1, we sought to validate this claim by presenting participants with individual comic panels containing either an Agent or Patient from a preparatory action and asking them to describe the content of the panel. If the representation of a preparatory action sufficiently allows for participants to identify an entity's semantic role, this would predict that participants would generate more "active" descriptions for panels with Agents in preparatory states than panels with Patients in preparatory states.

2.1. Methods

2.1.1. Participants

Fifteen comic readers (11 male, 4 female, mean age: 27.1) participated in the study. They were recruited through a link on the first author's blog, and were compensated with entrance into a raffle. Self-defined "comic readers" were chosen to ensure that participants were familiar with the materials and this manner of assimilating sequential pictures. Studies have shown comprehension differs based on comic reading ability (Cohn et al., 2012; Nakazawa, 2005), so fluent comic readers were used to reduce the heterogeneity in the population. Participants' comic reading fluency was assessed using a pretest questionnaire that asked them to self-rate the frequency with which they read various types of visual narratives (comic books, comic strips, graphic novels, Japanese comics, etc.), read books for pleasure, watched movies, and drew comics, both currently and while growing up. These ratings were measured using a scale of 1–7 (1 = never, 7 = always), and the questionnaire also gauged their self-assessed "expertise" at reading and drawing comics along a five-point scale (1 = below average, 5 = above average). A "fluency rating" was then computed using the following formula:

$$(Mean\ comic\ Reading\ Freq \times Comic\ reading\ expertise) + \left(\frac{Comic\ Drawing\ Freq \cdot \times Drawing\ Ability}{2}\right) \quad (1)$$

This formula weighted fluency towards comic reading comprehension, giving an additional "bonus" for fluency in comic production. Previous research has shown that the fluency score resulting from this metric provides a strong predictor of both behavioral and neurophysiological effects in online comprehension of visual narratives (Cohn et al., 2012). An idealized average along this metric would be a score of 12, with low being below 7 and high above 20. Participants' fluency was very high with a mean score of 26.65 (SD = 8.5).

2.1.2. Stimuli

We created 30 visual narratives using panels from *The Complete Peanuts* volumes 1–6 (1950–1962) by Charles Schulz. These sequences were either 6-panel long strips created by combining panels from disparate sequences into novel orders, or were based on existing Sunday strips varying in length from 7- to 12-panels long. To reduce the influence of written language, we only used panels without text or altered them to be without text. *Peanuts* was chosen because (1) panel sizes are consistent, with (2) regular characters and situations, which are (3) familiar to most readers, and (4) could be drawn from a large corpus.

From these whole strips (used as a whole in Experiment 4), we extracted 30 critical panels based on the following criteria: (1) the panel showed both entities, (2) the panel depicted the entities in a state just prior to a completed action that brought the two into contact, and (3) the Agent and Patient in this preparatory state were not in contact. In the panels chosen, Agent and Patient roles were associated with animate and inanimate entities. Animacy distribution of Agents was as follows: 18 humans, 10 animals, and 2 inanimate entities. For Patients, the distribution was: 10 humans, 12 animals, and 8 inanimate entities. Since previous work has shown directionality of action can correlate prototypically with semantic roles (e.g., Chatterjee, 2001, 2010), we also coded each panel for directionality of intended action. Of the 30 panels, 15 depicted left-to-right flow, 10 depicted right-to-left flow, and in the remaining five, flow was centered or lacked clear direction.

Using Adobe Photoshop, we divided each of the original panels into two panels, each featuring only the Agent or the Patient (as in Fig. 1). Panels were rescaled and cropped appropriately to maintain consistency of size and amount of panel filled by the entity. Additional examples of sample stimuli from all experiments, along with verbal descriptions of all critical panels, can be viewed in the Appendix.

2.1.3. Procedure and data analysis

We randomly ordered the 60 individual panels depicting Agents and Patients and presented them to participants in an online survey (surveymonkey.com). Participants were asked to describe the panels in a single, simple sentence.

The sentences were analyzed as representing either active or passive states. Each sentence was given either a "1" for active events or a "0" for passive states. Active events were considered as verbs

describing overt actions, like "reaching" and "winding up," as well as sentences clearly predicting a following event ("preparing to...", "is about to..."). Passive events consisted of states of being ("is happy/sad/etc.", no verb) or describing a passive process ("sitting", "looking", "thinking"). Three independent coders analyzed the data, and disagreements between assessments were resolved by choosing the coding agreed upon by 2 out of 3 coders. There was full interrater agreement for 44% of items, with the remaining 56% being resolved through majority opinion (i.e., agreement between 2 out of 3 raters). We then averaged these ratings across participants for each item, and the averages for Agent and Patient panels were then analyzed using a *t*-test. Because we wanted to measure the "activity" for each item, rather than aggregate "activity" across all times, only an items analysis was conducted.

Additional *t*-tests compared the panels in terms of animacy (animate vs. inanimate) and directionality of motion (left-to-right, right-to-left, excluding centered figures), by removing all scenarios with inanimate entities (resulting in 20 scenarios), or those with a canonical left/Agent-right/Patient (resulting in 22 scenarios).

2.2. Results and discussion

As depicted in Fig. 2, we found that Agent panels were described as significantly more active than Patient panels, $t(29) = 5.02$, $p < 0.001$. An additional analysis eliminating all scenarios featuring inanimate entities, thus leaving only animate ones, found that Agent panels remained more actively described than Patient panels, (Agents: 0.74 (0.31); Patients: 0.30 (0.41); $t(19) = 3.74$, $p < 0.005$). Similarly, analysis using only panels with a right/Agent-left/Patient (i.e. non-canonical) flow of action likewise found similar effects, (Agents: 0.75 (0.32); Patients: 0.19 (0.27); $t(21) = 4.45$, $p < 0.001$). The results are thus consistent with our proposal that participants readily establish semantic roles of entities in preparatory states prior to the presentation of an action.

3. Experiment 2: Prediction of upcoming event

The results of Experiment 1 suggest that participants do indeed perceive Agents in preparatory states as more active, thus more Agent-like, than Patients in preparatory states. In Experiment 2 we next sought to test whether participants were establishing the entities as Agents of *upcoming* actions, rather than simply Agents of preparatory actions.

Previous research has shown that comprehenders' ability to anticipate subsequent events depends largely on the constraints allowed by the context. Early research on discourse processing suggested that prediction is fairly poor for future narrative events on the level of plotlines (Magliano, Baggett, Johnson, & Graesser, 1993; Potts, Keenan, & Golding, 1988) or only occurs at a general level (McKoon
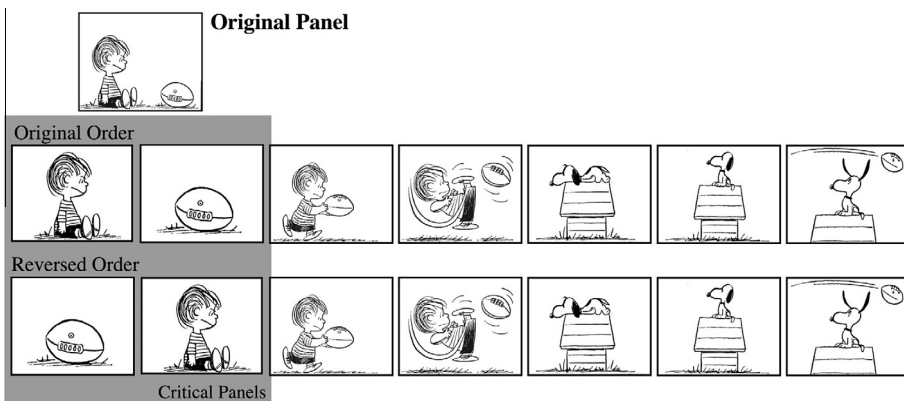

Fig. 3. Panels depicting passive actions where characters have no distinguishable semantic roles, which become divided into panels keeping each character distinct. Sequences can then order either character first (Experiment 3).

& Ratcliff, 1986). Subsequent studies found some evidence for prediction in both written (Fincher-Kiefer, 1993, 1996; Fincher-Kiefer & D'Agostino, 2004; van den Broek, Rapp, & Kendeou, 2005; Weingartner, Guzmán, Levine, & Klin, 2003) and filmed narratives (Magliano, Dijkstra, & Zwaan, 1996; Tan & Diteweg, 1996). Importantly, recent work by Szewczyk and Schriefers (2013) showed that readers can make predictions for coarse-grained semantic features even when narrative contextual constraint is relatively low (average cloze less than 0.32). Additionally, research on event comprehension has consistently shown that people can also predict real-world events somewhat prior to their actual occurrence (Lasher, 1981; Sitnikova, Holcomb, & Kuperberg, 2008; Zacks, Kurby, Eisenberg, & Haroutunian, 2011; Zacks, Speer, Swallow, Braver, & Reynolds, 2007; Zacks et al., 2001). Making such predictions benefits comprehension because the event has already begun to be built, thereby decreasing the subsequent cost of processing the actual events (Kurby & Zacks, 2008; Zacks et al., 2007). Together, these findings suggest that participants generate at least a limited set of predictions at the temporally local level.

Within our framework, such local event predictions may be motivated by the preparatory actions of Agents, rather than Patients. However, the previous research on real-world events used films with continuous representations of actions, meaning that preparatory and completed actions blend into each other. By using the discrete segmentation offered by visual narratives, we can more clearly address this role of prediction in visually depicted events. Thus, to determine whether preparatory Agents anticipate upcoming actions, we once again presented participants with the same panels depicting an Agent or Patient in a preparatory state. However, rather than asking participants to describe what was happening in the depicted panel, we instead asked, "What happens next?" If preparatory Agents are viewed as taking part in an *upcoming* action, participants should be able to generate consistent predictions about the upcoming action. Based on previous research, we would expect participants to make coarse-grained predictions regarding an immediately upcoming action based on Agents in preparatory states. For example, Lucy with her fist pulled back would generate predictions that she was about to punch (or punch at) *something* or *someone* (i.e., Linus). On the other hand, Patients in preparatory states would not be expected to generate specific predictions about upcoming events—Linus looking up in the air does not anticipate any particular subsequent actions. While the Patients are equally "temporally" close to the event about to take place as the Agents, they provide less constraint on the scope of actions that might occur, thus providing limited motivation for participants to generate strong predictions regarding upcoming events.

3.1. Methods

3.1.1. Participants

Twenty-four comic readers (20 male, 4 female, mean age: 28.9) participated in the study, and were recruited through a link on the first author's blog and compensated with entrance into a raffle. Participants were again rated as highly fluent in comic reading expertise, with an average score of 27.57 (SD = 9.86).

3.1.2. Stimuli

The same stimuli were used in Experiment 2 as in Experiment 1.

3.1.3. Procedure and data analysis

As in Experiment 1, the 60 individual panels depicting Agents and Patients were presented in a random order to participants in an online survey (surveymonkey.com). Using a single, simple sentence, participants were asked to describe what they thought would happen next by looking at the panel.

The sentences were analyzed as describing either predictable or non-predictable actions. A participant's description was deemed as "predictable" if it agreed with the outcome predicted by at least one other participant. For each participant, each item's description sharing a common interpretation was given a "1." In contrast, a score of "0" was given to descriptions that had no consensus, retained expectation of the status quo (i.e., "nothing happens" or predicted the same as the depicted content), or made no action-based predictions (i.e., predicted a non-sequitur about a new character). The same three coders were used as in Experiment 1. Again, full interrater agreement was 36% across all items,

with the remaining 64% being resolved through majority opinion (i.e., agreement between 2 out of 3 raters). We then generated a score of "predictability" for each item by averaging all scores for an item across participants. These averaged ratings for Agent and Patient panels were compared using a *t*-test. Because item predictability scores were based on averages generated across participants, only an items analysis was performed. Again, additional analyses excluded all scenarios with inanimate entities or canonical left/Agent-right/Patient directionality.

3.2. Results and discussion

Consistent with our hypothesis, Agents in preparatory states tended to elicit more consistent predictions about an upcoming action than did Patients in preparatory states, $t(29) = 3.06$, $p < .006$, as in Fig. 2. This pattern of results was also significant when only non-canonical right/Agent-left/Patient panels were analyzed, (Agents: 0.56 (0.28); Patients: 0.38 (0.09), $t(21) = 2.31$, $p < 0.05$). A similar, though only marginally significant effect was found when all scenarios with inanimate entities were eliminated, (Agents: 0.55 (0.22); Patients: 0.41 (0.17), $t(19) = 1.88$, $p < 0.08$).

These results suggest that Agents in preparatory states are indeed perceived as Agents of an upcoming event, not simply the Agent of an isolated preparatory action. Notably, as suggested by previous research on discourse, these predictions related to local subsequent actions—the immediate event after that panel—not to broader predictions at the level of plotlines. Importantly, this finding also suggests that viewing Agents in preparatory states may initiate the building of the overall event representation, prior to the central action of that event taking place (e.g., Zacks et al., 2001, 2007), consistent with the proposal that Agents play a central role in building event structure.

4. Experiment 3: Effects of sequential ordering prior to semantic role assignment

We next wanted to determine whether the re-ordering of characters from their original presentation would affect participants' online reading of comic strips. Looking at Fig. 3, we see that the original panel of Linus looking at a football can be split into two panels, one of Linus, one of the football, similarly to how our critical panel was split for Experiments 1 and 2 (see Fig. 1). In both cases, the two characters can then be presented in the same sequential order as in the original panel (assuming left-to-right reading), or transposed from this original sequential relationship. However, there is a crucial difference between the two cases. In Fig. 1, Lucy and Linus are in preparatory states, readily identifiable as the Agent and Patient of the subsequent action. Contrast this with Fig. 3, in which neither Linus nor the football can be meaningfully assigned semantic roles. According to our hypothesis, the sequential presentation order should affect processing when the semantic roles of the characters are already established (as in Fig. 1), but not when the semantic roles remain unspecified (as in Fig. 3). Previous research has shown that alteration of the left–right composition of images has little affect on comprehension of visual narratives (Gernsbacher, 1985), though experimental evidence testing this reversal with individuated characters has yet to be done.

As our critical manipulation in Experiment 4, the ordering of semantic roles within a sequential narrative, required characters to be presented both in their original order and in reverse order, in Experiment 3 we sought to determine whether deviation from original sequential presentation, in and of itself, would affect online visual narrative comprehension in a self-paced reading paradigm. Based on our hypothesis, we would expect that viewing times would be the same for "Original Order" as "Reversed Order" (Fig. 3) strips. On the other hand, if low-level visual features somehow provide readers cues as to original order, this would predict that participants should read panels in the "Original Order" sequences, faster than those in the "Reversed Order" sequences.

As a comparison condition, we also included "scrambled" sequences, which lacked coherence and thus preventing the build-up of narrative structure. One the whole, comprehenders "lay a foundation" of meaning early in the sequence and "build structure" as they progress through a coherent sequence (Gernsbacher, 1985). This buildup of structure consistently across the course of a sequence, whether in sentences (Marslen-Wilson & Tyler, 1980; Van Petten & Kutas, 1991), verbal discourse (Glanzer, Fischer, & Dorfman, 1984; Greeno & Noreen, 1974; Haberlandt, 1984; Kieras, 1978; Mandler &

Goodman, 1982), or visual narratives (Cohn et al., 2012; Gernsbacher, 1985). Because this process of buildup requires a coherent grammatical structure, it is more difficult in sequences where the order of units is scrambled, such as the rearrangement of words in sentences (Marslen-Wilson & Tyler, 1980; Van Petten & Kutas, 1991), or a scrambled string of panels in a visual narrative (Cohn et al., 2012; Gernsbacher, Varner, & Faust, 1990). Thus our comparison condition served both to replicate our previous findings as well as ensure participants stayed on task.

4.1. Methods

4.1.1. Participants
Twenty-eight experienced comic readers (19 male, 9 female, mean age: 21.04) from the Tufts University undergraduate population participated in the experiment for compensation. All participants gave their informed written consent according to Tufts University Human Subjects Review Board guidelines. Participants' fluency, $M = 15.13$ (SD = 8.48), near the idealized average of 12.

4.1.2. Stimuli
We created 30 visual narratives using panels from *The Complete Peanuts* volumes 1–6 (1950–1962) by Charles Schulz. Critical panels from positions 1–5 were divided into two separate panels using Adobe Photoshop, so that they featured only a single character. Critical panels were determined using criteria that included the semantic properties of: (1) the introduction of a referential relationship, and/or (2) characters in a passive state of being (sitting, standing, etc.), or (3) characters in the midst of an unbounded process (such as walking).

Critical panels were coded based on whether elements appeared on the left or the right side of the original single panel. Divided panels were ordered to create two alternative sequence types: Original Order or Reversed Order, as depicted in Fig. 3. These two sequence types were then counterbalanced into two lists, such that no list repeated strips.

Forty filler strips were created that matched normal strips with a version that scrambled the order of panels. Critical panels in the normal strips matched the position in their corresponding scrambled strips. To each of the experimental lists we added 20 normal and 20 anomalous filler strips, counterbalanced so that within each list, a given filler strip appeared only once, and across both lists, each filler strip appeared in the normal and anomalous version.

4.1.3. Procedure
Participants controlled the pace of viewing each strip frame-by-frame on a computer screen. Viewing times were measured to each button press for how long each frame stayed on screen. Trials began with a black screen with the word "READY" in white, followed by a fixation cross (+), and then each panel centered on an otherwise black screen with a 300 ms ISI. After each sequence, a question mark (?) appeared, where participants rated how easy the whole strip was to understand (1 = difficult, 7 = easy). A practice list with ten stimuli preceded the actual experiment to orient participants to the procedure.

4.1.4. Data analysis
A 2 (critical panel type: left vs. right) × 2 (critical panel order: original vs. reversed) × 5 (sequence position: panels 1/2, 2/3, 3/4, 4/5, 5/6) ANOVA was used to compare viewing times of critical panels. Interactions involving sequence position were followed up with polynomial contrasts across sequence position. Interactions involving critical panel type and critical panel order were collapsed across sequence position and analyzed with a 2 (critical panel type) × 2 (critical panel order) ANOVA. Because we did not have an equal number of stimuli for each ordinal position, our analyses involving sequence position were undertaken only using a subjects analysis. However, all other analyses used both a subjects and items analysis.

4.2. Results

4.2.1. Ratings

Experimental sequences were rated as easily understandable with no difference in ratings between Original Order (M = 5.76, SD = 1.04) and Reversed Order sequences (M = 5.58, SD = 0.98), $t(1,27)$ = 0.84, p = 0.41.

4.2.2. Viewing times

We found main effects of critical panel order, $F(1,27)$ = 13.35, $p < 0.005$, and sequence position, $F(4,108)$ = 10.05, $p < 0.001$. However, we found no main effects of critical panel type, $F(1,27)$ = 0.305; p = 0.585. See Fig. 4. Thus the viewing time for critical panels was only affected by how far into the sequence it was found, not by the relative ordering.

Significant interactions were found only between critical panel order and sequence position, $F(4,108)$ = 17.76, $p < .001$, reflecting that viewing times were far slower at the start of sequences than at subsequent positions. Follow up polynomial contrasts found significant linear, quadratic, and cubic trends for ordinal position (all $Fs > 8.8$, all $ps < .01$) and an interaction between critical panel position and ordinal position (all $Fs > 4.5$, all $ps < .05$).

To further confirm that the alternation between left–right orders did not affect viewing times, we analyzed critical panels by collapsing across sequence position (see Fig. 5). Again, main effects were found for critical panel order, $F1(3,81)$ = 17.01, $p < .001$, $F2(1,29)$ = 9.69, $p < .005$, but not critical panel type, $F1(1,27)$ = .372, p = .631, $F2(1,29)$ = .002, p = .962. There was also no interaction between critical panel order and critical panel type, $F1(3,81)$ = .833, p = .480, $F2(1,29)$ = 2.04, p = .164.

A second analysis examined only critical panels containing animate characters, eliminating all experimental scenarios with inanimate entities. No main effects were found for critical panel type, $F1(1,27)$ = 1.97, p = .172, $F2(1,14)$ = .647, p = .435, nor was the interaction between critical panel type and critical panel order significant, $F1(1,27)$ = 3.07, p = .091, $F2(1,14)$ = 2.17, p = 163. However, there was a significant main effect for critical panel order, $F1(1,27)$ = 15.66, $p < .001$, $F2(1,14)$ = 6.20, $p < .05$.

Finally, we sought to confirm that our null results were not due to lack of statistical power or due to participants not paying attention to the task by examining the viewing times to critical panels in normal sequences compared to those in scrambled sequences. Critical panels in scrambled sequences
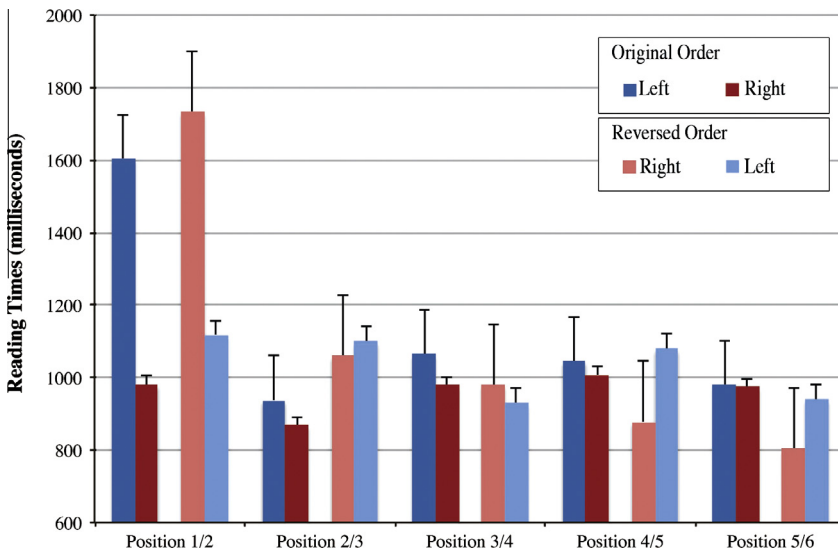

Fig. 4. Mean viewing times for critical panels that had no distinct semantic roles across sequence positions (Experiment 3).
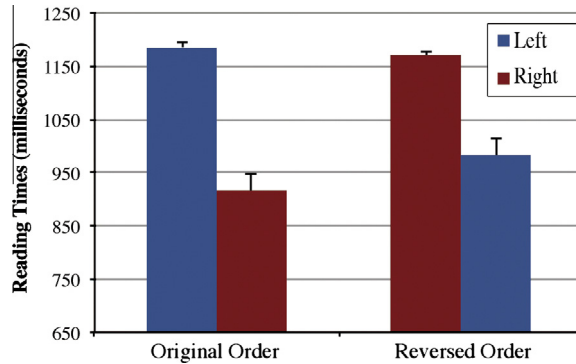
Fig. 5. Mean viewing times for critical panels of characters with no distinct semantic roles based on the ordering of characters, collapsing across sequence position (Experiment 3).

with randomly ordered panels took significantly more time to view than critical panels in normal sequences, $t1(27) = -2.25$, $p < .05$, $t2(58) = -2.01$, $p < .05$.

4.3. Discussion

This experiment sought to determine whether displacing characters from their original orientation within a visual narrative sequence, prior to them being assigned semantic roles, would affect viewing times. Overall, our hypothesis was borne out: we found that the order of presentation failed to modulate viewing times on critical panels. However, viewing times for initial panels were longer compared to panels later in the sequence, as evidenced by a significant main effect of sequence position. The effect was most likely driven by the "laying a foundation" of the narrative structure suggested by the slower viewing times at the first panel position, consistent with previous studies (Cohn et al., 2012; Gernsbacher, 1985). In the first position, characters are introduced for the first time, and are thus viewed slower, while at positions after this, characters have already been established and thus do not require as much time to comprehend. These initial panels thus establish referential entities of the sequence prior to their actions in the subsequent panels.

5. Experiment 4: Effects of semantic role ordering

In Experiments 1 and 2, we demonstrated that participants could reliability differentiate Agents from Patients for an upcoming action when viewing panels containing entities in their preparatory states. In Experiment 3 we demonstrated that serial order of individual entities within visual narratives does not, in and of itself, alter viewing times for the panels.

Having validated our methodological assumptions, we next turned to addressing the central question of our inquiry: how does the sequential ordering of semantic roles influence their processing? Recall from the Introduction that previous studies have found an Agent advantage under a variety of experimental paradigms. Within language-based studies, Agent First syntactic structures have been shown to be more frequent cross-linguistically (Kemmerer, 2012) and to be understood more easily (Ferreira, 2003). Previous studies have also demonstrated that when communicating manually, Agents tend to precede Patients (Goldin-Meadow et al., 2008). Similarly, participants tend to prefer Agents ahead of Patients when ordering or viewing pictures (Gershoff-Stowe & Goldin-Meadow, 2002), and tend to more quickly identify Agents than Patients both in pictures (Segalowitz, 1982) and films (Germeys & d'Ydewalle, 2007).

We have proposed that all of these findings can be explained within a single framework in which the Agent serves to initiate the generation of event structure. The Agent provides the proximal cause of an action taking place, thus attending to the Agent allows one to orient to what is *about* to happen.

From this perspective, argument ordering within linguistic structures could be a consequence of this attention focusing. Indeed, it is generally accepted that the initial argument of a sentence makes the referent the focus of attention (Gernsbacher & Hargreaves, 1988). Focusing attention on the Agent could have advantages because, as shown in Experiment 2, Agents are associated with more predictions about the upcoming action than Patients.

In Experiment 4 we tested this hypothesis by manipulating the serial order of semantic roles within visual narratives. If Agents play a more central role in the initial building of event structure than do Patients, viewing times for Agents should be longer than viewing times for Patients. Additionally, we would expect this effect to be modulated by serial order. Agents in Agent First (AF) ordering would be expected to incur longer viewing times than Agents in Patient First (PF) orderings. While in both cases the Agent would serve to initiate event building processes, in AF-orderings, Patients should easily be incorporated into the established event structure. In PF-orderings, the Patient would not provide enough information to build the event, which would be needed by the subsequent Agent panel.

Alternatively, if the canonical Agent First ordering observed in language and gesture is not epiphenomenal, as we have argued, but instead is due to a canonical Agent–Patient schema, this would predict that both types of panels should be viewed longer when the schema is violated (i.e., Patient followed by Agent) than when it is not. Our null hypothesis was that, as in Experiment 3, serial ordering of semantic roles would have no effects beyond the ordinal position, with viewing times being longer for the first panel compared to the second, and that this effect would not be modulated by semantic role ordering.

5.1. Methods

5.1.1. Participants

Sixty Tufts University undergraduates (33 male, 27 female, mean age: 20.16) experienced in reading comics participated in the experiment for compensation. All participants gave their informed written consent according to Tufts University Human Subjects Review Board guidelines. Data from four participants was excluded because they did not appear to understand the task. Participants had an average comic reading fluency score of 12.05 (7.36).

5.1.2. Stimuli

We used the full sequences discussed in the Stimuli for Experiment 1. Here, the divided panels (Agent, Patient) were used to create two sequence types: Agent First (AF) and Patient First (PF) orders, as in Fig. 1. The ordinal position of critical panels in the sequence was varied throughout strips depending on their original location, and never appeared at the sequence final position. This yielded three manipulated variables: critical panel position (first or second), semantic role (Agent or Patient), and ordinal sequence position (positions 1/2–6/7).

The two experimental sequence types were randomly distributed throughout two lists using a Latin Square Design such that each participant never saw the same strip twice. In addition to these experimental sequences, we included 40 fillers with varied degrees of coherence.

5.1.3. Procedure

The same procedure was used in Experiment 4 as in Experiment 3.

5.1.4. Data analysis

Ratings of strips were compared using a *t*-test. Viewing times to button presses for each panel were analyzed using an analysis of variance (ANOVA) in a 2 (semantic role: Agent vs. Patient) × 2 (critical panel position: first vs. second) × 6 (sequence position: panels 1/2, 2/3, 3/4, 4/5, 5/6, 6/7) design to compare semantic roles across various positions in the sequence for the subjects analysis. Follow ups to interactions between semantic role and sequence position collapsed across critical panel position and compared Agents and Patients at each sequence position using paired *t*-tests. Additionally, interactions involving sequence position were also followed up with analyses of polynomial trends across positions for each critical panel's semantic role individually, using regressions that set sequence position as the predictor. Because we did not have an equal number of stimuli for each ordinal

position, our analyses involving sequence position used only a subjects analysis. Interactions between critical panel position and semantic role were followed up with additional ANOVAs at the critical panel position for both subjects and items analyses.

In order to ensure that effects could not be attributed to the difference between animate and inanimate entities, a second analysis eliminated all scenarios featuring inanimate entities, leaving data from 20 sequence types. These critical panels were collapsed across panel position and analyzed using a 2 (semantic role) × 2 (critical panel position) ANOVA. Follow up analyses again used *t*-tests to compare individual interactions.

5.2. Results

5.2.1. Ratings

Participants rated both types of sequences as easily understandable (Agent First orders: 6.18 (0.48); Patient First orders: 6.24 (0.47)), and did not differ from each other, $t(1,55) = 1.4$, $p = 0.16$.

5.2.2. Viewing times

We first compared viewing times for semantic roles and critical panel positions across ordinal sequence position, depicted in Fig. 6. We found main effects for critical panel position, $F(1,55) = 18.49$, $p < 0.001$, and sequence position, $F(5,275) = 40.4$, $p < 0.001$, but not for semantic role, $F(1,55) = 0.337$; $p = 0.564$. We also found two-way interactions between semantic role and critical panel position, $F(1,55) = 102.63$, $p < 0.001$, and critical panel position and sequence position, $F(5,275) = 9.29$, $p < 0.001$, and a three-way interaction between all factors: semantic role, critical panel position, and sequence position, $F(5,275) = 40.9$, $p < 0.001$.

The interaction between semantic role, critical panel position, and sequence position in the ANOVA reflected that, overall, viewing times decreased across ordinal position in the sequence based on the position of each semantic role. Polynomial contrasts found significant linear, quadratic, and cubic trends for ordinal position (all $Fs > 4.5$, all $ps < 0.05$), an interaction between critical panel position and ordinal position (all $Fs > 7.7$, all $ps < 0.01$), and an interaction between semantic role, critical panel position, and ordinal position (all $Fs > 4.8$, all $ps < 0.05$).
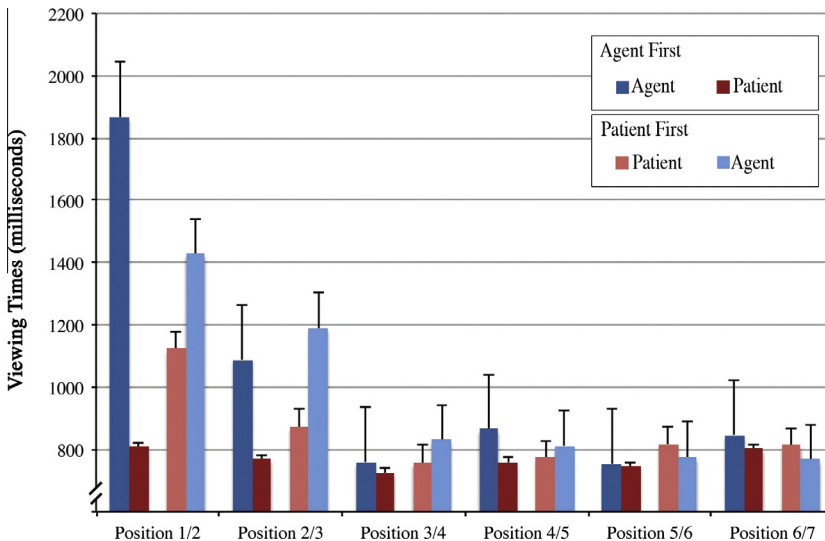

Fig. 6. Mean viewing times for reversal in the order of Agents and Patients in critical panels across ordinal sequence position (Experiment 4).

We followed up the interaction between semantic role and sequence position by comparing semantic roles (Agent vs. Patient) at each ordinal position of the sequence. Collapsing across critical panel position, we found that Agents were viewed longer than Patients at each sequence position, with the greatest difference arising earlier in the sequence. Viewing times between semantic roles differed at positions 1/2, 2/3, and 4/5 (all $ts > 2.4$, all $ps < 0.05$), but did not differ at positions 3/4, 5/6, and 6/7 (all $ts < 1.6$, all $ps > 0.107$).

Next, we followed up the interaction between semantic role and sequence position by analyzing the trends of viewing times of individual semantic roles across sequence position. This analysis used regressions, placing sequence position as the predictor in order to assess the trends of viewing times for panels across sequence position. Linear effects appeared for all critical panels (all $Fs > 8.8$, all $ps < 0.005$), except AF-Patients, $F(1, 334) = 0.022$, $p = 0.88$, summarized in Table 1. Viewing times to AF-Agents grew faster the most dramatically across sequence position, followed by PF-Agents, and PF-Patients. AF-Patients were viewed at a constant pace across positions.

Finally, we compared the interaction between critical panel position and semantic role by collapsing across sequence position. Our subjects analysis found main effects for critical panel position, $F1(1, 55) = 11.20$, $p < 0.005$, $F2(1, 29) = 2.77$, $p = 0.11$, but no main effects were found for semantic role, $F1(1, 55) = 0.632$, $p = 0.430$, $F2(1, 29) = 1.84$, $p = 0.186$. Interactions appeared between critical panel position and semantic role in all analyses, $F1(1, 55) = 30.92$, $p < 0.001$, $F2(1, 29) = 11.16$, $p < 0.005$. As depicted in Fig. 7, AF-Agents had the longest viewing times, while AF-Patients had the shortest viewing time, with viewing time for PF-Patients and PF-Agents falling in between, though both having shorter viewing times than AF-Agents.

A second analysis examined only critical panels containing animate characters, eliminating all experimental scenarios with inanimate entities. Significant main effects were found in the subjects analysis for semantic role, $F1(1, 55) = 20.97$, $p < 0.001$, $F2(1, 19) = 0.944$, $p = 0.344$, but no significant main effects were found for critical panel position, $F1(1, 55) = 0.024$, $p = 0.876$, $F2(1, 19) = .335$, $p = 0.569$. There was also a significant interaction between semantic role and critical panel position, $F1(1, 55) = 4.21$, $p < 0.05$, $F2(1, 19) = 3.86$, $p = 0.064$. The same overall pattern of results appeared: Agents were always viewed longer than Patients. In particular, AF-Patients were viewed the longest, 909 (290), followed by PF-Agents, 897 (181), PF-Patients, 803 (180) and finally AF-Patients, 750 (181). These viewing times to only animate panels appeared to be longer than those with inanimates, possibly due to animates being more visually complex than inanimates (i.e., animates, like people or animals are more decomposable into meaningful elements, such as "head" and "body," than inanimate entities, like footballs or snowballs).

5.3. Discussion

The results of Experiment 4 clearly demonstrate that participants were sensitive to both semantic roles assigned to entities in individual panels and the order in which the semantic roles appeared. On the whole, viewing times for Agent panels were longer than for Patient panels and this difference was significantly greater when Agent panels appeared prior to Patient panels. Across sequence position, viewing times for Agents were longest at the beginning of a sequence and rapidly decreased across sequence position. This pattern was most pronounced for Agents appearing before Patients. Interestingly, while a similar pattern, though with shallower slope, was observed for Patients appearing before Agents, when Patients followed Agents, their viewing times were unaffected by sequence position.

Table 1
Linear regressions for critical panels across panel positions.

	Agent First		Patient First	
	Agent	Patient	Patient	Agent
$r =$	0.169[*]	0.000	0.026[*]	0.127[*]
$b =$	−171.43	−1.8	−48.70	−130.55

r = Correlation coefficient, b = slope in milliseconds ($n = 6$).
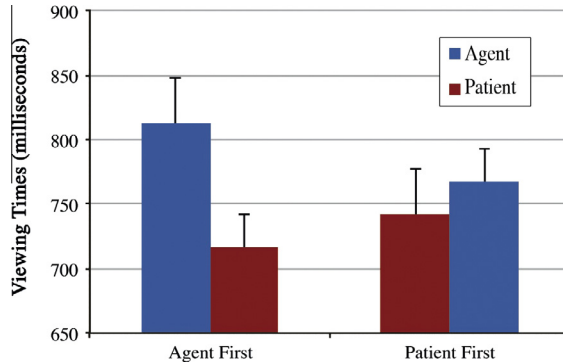[*] $p < .005$.

Fig. 7. Mean viewing times for critical panels with alternate orders of Agent and Patients, collapsing across sequence position (Experiment 4).

Our findings suggest that participants did not use a canonical sequential ordering schema while reading the visual narratives. If participants had relied on this type of schema, viewing times for both Agent and Patient panels would have been shorter when the two roles occur in the canonical Agent First order compared to when they appeared in reverse, Patient First, order. Instead, we found that viewing times for Agents were actually reduced when they followed Patients compared to when they preceded Patients, the exact opposite of what would be expected if participants relied on a canonical Agent–Patient schema to guide their comprehension.

Our findings are consistent with the proposal that Agents initiate the building of event structure and the generation of predictions about the upcoming event. First, viewing times for Agents were longer, relative to Patients, and this slowing was most pronounced when Agents preceded Patients. However, these slower viewing times to Agents appeared independent of semantic role ordering early in the sequence (positions 1/2, 2/3, and 4/5), when all viewing times were longer than those later in the sequence. Second, viewing times for Agents were significantly reduced when they appeared after Patients. This finding is consistent with the idea that Agents trigger predictions about an upcoming event. When a Patient had already been presented, predictions should be easier to generate as participants had more information about the upcoming action and no predictions regarding who or what will be Patient of the upcoming action.

The trends of viewing times across ordinal sequence positions further support this hypothesis. Overall, the decrease in viewing times across sequence positions is consistent with previous research showing that structure builds up across the course of a verbal discourse (Glanzer et al., 1984; Greeno & Noreen, 1974; Haberlandt, 1984; Kieras, 1978; Mandler & Goodman, 1982) or visual narrative (Cohn et al., 2012; Gernsbacher, 1985). However, viewing times for Agents decreased the most across sequence position, further supporting that they enable predictions about subsequent events. If slower viewing times for Agents reflect the cost of such predictions, as panels approach the end of a sequence prediction would be less necessary. Our results support such a conclusion: the greatest difference in viewing times between semantic roles occurred at the start of the sequence, while these differences diminished towards positions later in the sequence. Additionally, across sequence positions, viewing times stayed constant and short for Patients coming after Agents. We argue that, because the Agent had already established the event structure into which the Patient could be integrated, viewing times to subsequent Patients remained constant across ordinal panel positions, most likely at a ceiling. Importantly, Patients that could not yet be integrated into an established event structure (i.e., those appearing prior to Agents within a sequence), were viewed longer and the viewing times for these Patient panels decreased monotonically across sequence position, similarly to the pattern we observed for Agents.

Based on our findings in Experiment 3, the results of Experiment 4 are unlikely to be due to displacement of characters appearing in an order different from their original panel. Recall that viewing

times for panels with unspecified semantic roles were affected only by ordinal position, but not by reversal of presentation order (right before left). Our results also cannot be explained by any properties of Agent panels separate from their semantic role assignment, as we found a similar pattern of results when we controlled for animacy or left/right orientation within the original panel.

6. Experiment 5: Influence of semantic roles on event processing

Both the results of Experiments 2 and 4 are consistent with the idea that Agents, but not Patients, in preparatory states lead participants to predict an upcoming action and potentially initiate the building of event structure. In Experiment 5 we tested this hypothesis more directly by examining the influence of Agents and Patients in preparatory states on the subsequent processing of the actual event depiction within visual narratives. Specifically, in Experiment 5 we sought to determine whether completed actions are processed more quickly when preceded by Agents in preparatory states compared to Patients in preparatory states.

We have argued that comprehenders utilize Agents to initiate the building of event structure representation and that Agents in preparatory states lead participants to predict the immediately upcoming action in which the Agent will take part (as shown in Experiment 3). On the one hand, this leads to a greater initial processing cost in association with viewing Agents vs. Patients in preparatory states (as shown in Experiment 4). On the other hand, we suggest that this frontloading of processing should lead to faster processing of the subsequently presented completed action.

Once again, we presented participants with short sequences in which the central action involved an Agent and Patient interacting. In contrast to Experiment 4, however, within a given strip participants only viewed the Agent or Patient in a preparatory state just prior to viewing the subsequent action (see Fig. 8). We predicted that if Agents in preparatory states generate prediction and/or initiate the construction of an event representation, participants' viewing times for the subsequently presented event should be reduced when preceded by panels with Agents in preparatory states compared to Patients in preparatory states.

6.1. Methods

6.1.1. Participants

Twenty-four experienced comic readers (15 male, 9 female, mean age: 20.3) from the Tufts University undergraduate population participated in the experiment for compensation. All participants gave their informed written consent according to Tufts University Human Subjects Review Board guidelines. Participants had an average comic expertise rating of 13.83 (6.05), near the idealized average (12).

6.1.2. Stimuli

Thirty sequences were created using the same methods as in Experiments 1–4. Once again, panels with entities in preparatory states were divided into single panels, each containing either the Agent or
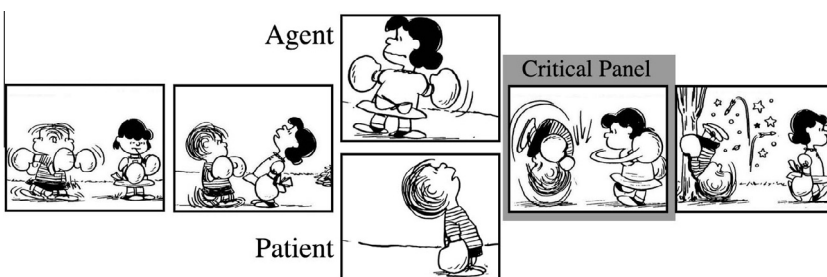

Fig. 8. Visual narrative sequences where panel depicting completed actions were preceded by panels showing either an Agent or a Patient.

Patient for the subsequently depicted action. To maximize our effects, we deliberately chose critical panels towards the start of panel sequences (positions 2, 3, and 4). Within these 30 sequences, 27 of Agents and 20 of Patients were animate. 14 Agents and 16 Patients featured a left-to-right flow (i.e., were facing right and/or were on the left side of the original undivided panel). Panels prior to the critical panel had equal amounts of the future-Agent on left side (13) and right side (13) of the panel, with 4 centered.

We created two versions of each strip where either the individuated Agent or Patient panel appeared in place of the original panel showing both entities. All critical panels depicting the completed action appeared in the panel immediately following the Agent or Patient panels (see Fig. 8). These sequences were then counterbalanced into two lists such that no list repeated strips, along with 60 filler sequences featuring varying degrees of coherence.

6.1.3. Procedure

The same procedure was used as in Experiments 3 and 4.

6.1.4. Data analysis

Viewing times were measured to button presses at the critical panel—showing the completed action—for each sequence type and compared using paired *t*-tests. As in Experiments 1–4, additional analyses eliminated all scenarios with inanimate entities and a canonical left/Agent-right/Patient orientation.

6.2. Results

Once again, experimental sequences were rated as easily understandable, with no difference in rating between the two sequence types in either subjects or items analysis (Agent-sequences: 5.91 (0.57); Patient-sequences: 5.94 (0.69); $ts < 0.38$, $ps > 0.71$). Viewing times for the critical panel were shorter following Agent than Patient panels (Fig. 9). This difference was significant in the subjects analysis, $t1(23) = 2.15$, $p < 0.05$, and approached significance in the items analysis, $t2(29) = 1.69$, $p = 0.10$. This pattern also appeared when excluding all inanimate entities (Action after Agents: 1264 (500); Action after Patients: 1501 (727), again only in the subjects analysis, $t1(23) = -2.79$, $p < 0.02$, $t2(16) = -1.14$, $p = 0.27$. This same pattern of results occurred when excluding all scenarios preceded by left positioned-Agents and right-positioned-Patients, (Action after Agents: 1310 (550); Action after Patients: 1103 (386), $t1(23) = -2.43$, $p < 0.03$, $t2(14) = -1.24$, $p = 0.23$.
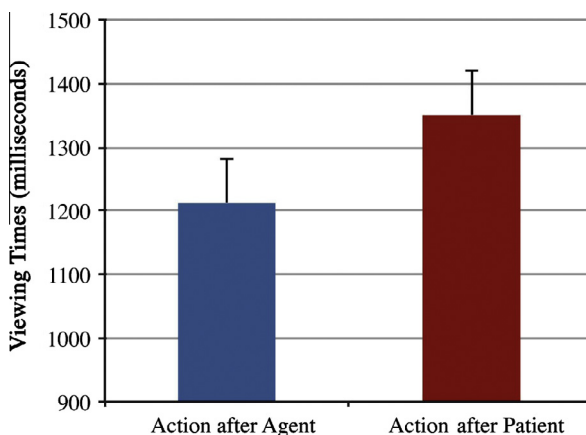

Fig. 9. Mean viewing times for panel depicting completed actions that were preceded by either a panel of an Agent or a Patient (Experiment 5).

Overall, these effects support that completed actions preceded by Agents are viewed faster than actions preceded by Patients. However, these effects only appeared to be significant in the subjects analysis. Because items analysis effects was only marginal, we sought additional support for the idea that Agents initiate the building of event structures by reanalyzing the panels depicting completed actions that followed our critical manipulations of serial order (Experiment 3) and semantic roles (Experiment 4). We analyzed the viewing times to panels depicting completed actions that immediately followed the critical alternation between Agent and Patient panels in Experiment 4. We hypothesized that viewing times for completed actions following an Agent First order would be shorter than those following a Patient First order, because the initial engagement with the Agent allows for easier building of the overall event structure. This was indeed the case. Viewing times for panels depicting completed actions following Agent First orders were shorter than those following Patient First orders (Agent First: 944 (182); Patient First 1050 (168); $t1(55) = -2.17$, $p < 0.05$, $t2(29) = -2.03$, $p = 0.052$). We then sought to confirm that this effect could not be due to serial order alone by examining completed actions in Experiment 3 preceded by alternating panels where the entities had no distinct semantic roles. Here, panels containing completed actions did not appear immediately after the critical manipulated panels, with varying numbers of intervening panels. We thus analyzed both panels with completed actions and those immediately following the critical manipulated panels. We found no difference between viewing times for Original Orders vs. Reverse Orders in either case, be they for panels with completed actions (Original Orders: 1140 (445); Reverse Orders: 1177 (433); $t1(27) = -.857$, $p = 0.399$, $t2(29) = -.800$, $p = 0.430$), or for panels that immediately followed the critical panels (Original Orders: 1102 (404); Reverse Orders: 1111 (388); $t1(27) = -.195$, $p = 0.847$, $t2(29) = -.197$, $p = 0.845$).

6.3. Discussion

These findings are consistent with the hypothesis that participants can use predictions generated in response to viewing Agents in preparatory states to facilitate the processing of immediate, local upcoming events. In Experiment 5, viewing times for panels depicting completed actions were shorter when preceded by an Agent in preparatory state relative to a Patient in preparatory state. We suggest that upon viewing Agents in preparatory states, participants were able to frontload the building of event structure as well as make predictions about the upcoming event, thus reducing the amount of processing needed at the point of presenting the completed action. When participants only saw the Patient in a preparatory state, event structure building could not commence until the Agent was presented at the subsequent panel (i.e., the depiction of the completed event). Thus in this case, participants incurred processing costs associated with event structure building at the completed event, leading to longer viewing times compared to when the completed event was preceded by an Agent.

Our findings are reinforced by the further analysis of completed actions in Experiments 3 and 4. The analysis of panels from Experiment 4 showed that panels depicting completed actions were viewed faster following semantic roles in their canonical Agent First order than when following the non-canonical Patient First order. We interpret these results as reflecting the easier access to the overall event structure offered by initial viewing of the Agent. When participants reached the Agent first, they were able to establish information about the whole event, leading to faster viewing of the subsequent Patient (as discussed in Experiment 4) and of the subsequent completed action. When viewing the Patient first, no overt event structure is immediately available, resulting in somewhat slower processing at the subsequent Agent (as discussed in Experiment 4), and significantly slower viewing of the subsequent completed action. These effects are consistent with what might be expected from theories postulating a canonical Agent–Patient-Act schema (e.g., Ferreira, 2003; Goldin-Meadow et al., 2008), though serial order alone cannot explain these effects. Viewing times did not differ for panels depicting completed actions that were preceded by entities in indistinct semantic roles with differing serial orders, as in the times from Experiment 3. Furthermore, viewing times were faster for completed actions that were preceded by single Agents than single Patients (Experiment 5), where no serial order between semantic roles was present at all. This provides evidence that preparatory Agents motivated this effect, not simply an Agent–Patient ordered schema.

Together, these findings support that Agents in preparatory actions allow a participant to access an event structure more than a Patient in a passive state. These findings are consistent with, though lend insight into the possible causes of, previous proposals that prediction affords more efficient processing of completed events in real-world contexts (e.g., Kurby & Zacks, 2008; Zacks et al., 2001, 2007, 2011).

7. General discussion

In the current investigation, we utilized a novel methodology to examine the proposal that the "Agent advantage" observed across diverse modalities is due to Agents initiating the building of event structure and/or predicting immediate upcoming actions. In Experiment 1, we demonstrated that within panels of visual narratives of sequential images, participants readily identify Agents as being more active than Patients in preparatory states, i.e., prior to the action being overtly depicted where such semantic roles are confirmed. In Experiment 2, we demonstrated that Agents in preparatory states elicit a greater level of prediction about upcoming actions than Patients in preparatory states. In Experiment 3, we demonstrated that the re-ordering of characters prior to assignment of semantic roles did not influence online processing of visual narratives. In Experiment 4, we demonstrated that Agents are predominantly viewed longer than Patients independent of serial order position, especially earlier in the sequence. Additionally, in Experiment 4, we demonstrated that this increased processing cost for Agents was attenuated when the Patient of the upcoming action did not have to be predicted. Finally, in Experiment 5 we demonstrated that completed actions preceded by preparatory Agents were viewed faster than those preceded by Patients. Further analysis of the earlier experiments then confirmed that completed actions preceded by Agent First orders were viewed faster than those preceded by Patient First orders, while no difference in viewing times was found to completed actions preceded by different serial orders of entities with indistinct semantic roles. Taken together our findings are consistent with the proposal that Agents initiate the building of event structure and/or are used online to predict upcoming actions. We discuss our results in relation to previous findings below.

7.1. Serial ordering of semantic roles

As discussed in the Introduction, there is a strong cross-linguistic preference for placing Agents before Patients as the canonical serial order of semantic roles. This preference is shown whether a language uses SVO constructions, such as English, or SOV constructions, such as German. Indeed, some have proposed that the Agent serves as the default semantic role assigned to grammatical subjects (Keenan, 1976). This Agent before Patient ordering has also been observed within the realm of manual communication. For example, the homesign systems used by deaf children not taught formal sign language consistently use Agent–Patient-Act orders (e.g., Goldin-Meadow, 2003; Goldin-Meadow & Feldman, 1977). This same Agent–Patient-Act pattern appears when speaking individuals communicate through gesture alone (Gershoff-Stowe & Goldin-Meadow, 2002) and appears to be independent of the language spoken by the participants (Goldin-Meadow et al., 2008). Finally, Agent–Patient-Act orders also appear when participants arrange images to reconstruct events (Goldin-Meadow et al., 2008).

It has been proposed that this pattern of Agent before Patient emerges due to the fact that the Agent heads the causal chain that affects the Patient (Kemmerer, 2012). As a result, the canonical serial order mirrors this chain of causality, placing the Agent at the head of the grammatical construction. Here we take this proposal one step further and suggest that the Agent is placed at the head of grammatical constructions because it initiates the construction of an event representation, thus facilitating overall event comprehension. Thus, Kemmerer's "subject salience principle" becomes more of an "Agent salience principle" whereby Agents provide more information about event structure than Patients.

Why, then, would grammatical structures in sentences placing the Patient prior to the Agent result in slower reading times and poorer comprehension? Recall that Agents have been proposed as a default semantic role assigned to grammatical subjects (Keenan, 1976), and within language it is impossible to unambiguously specify the semantic role of an argument prior to the verb. Taken together, this would suggest that participants first interpret any initial noun as the Agent of an upcoming verb. Within our proposal, this would result in the reader or listener generating predictions about the

upcoming action and begin the construction of an event representation based on these predictions. Upon reaching additional syntactic information that contradicts this initial interpretation, the reader or listener would have to: (1) generate a new event representation based on the newly encountered argument, and/or (2) inhibit the previous event representation. Either one or both of these would incur additional processing costs.

Support for this proposal comes from a more detailed look of the results of Ferreira (2003), who found that participants had the most trouble with passive sentences such as *The woman was visited by the man*, where Agents were only correctly identified 71% of the time. This contrasted with sentences such as *The cheese was eaten by the mouse* where accuracy in identifying the Agent reached a ceiling (98%). In the first case, a fully plausible event structure can be built based on interpreting the subject argument as the Agent, making the inhibition of that structure difficult. On the other hand, one would be hard pressed to generate a full event structure where *cheese* is an Agent, rendering these structures to be fairly easily inhibited. Similar results were reported by Traxler, Morris, and Seely (2002) for object relative clauses, another Patient before Agent construction. They found that the widely reported reading time penalty for object, vs. subject, relative clauses was attenuated when the initial argument was a poor Agent. Again, within the proposed framework, these arguments would have led to rather limited event representation, thus making them easier to inhibit when presented with the actual Agent argument.

7.2. The Agent advantage in serial order

As discussed in the Introduction, across diverse domains, participants have been shown to respond more quickly to Agents than Patients. In her investigation of semantic role ordering in sentences, Ferreira (2003) found that participants were faster to make decisions regarding Agents than Patients in both active and passive constructions. Segalowitz (1982) found that participants identify Agents faster than Patients in simple graphic representations of events (line drawings of one fish biting another). When viewing films, participants more quickly orient to Agents than Patients (Germeys & d'Ydewalle, 2007) and this phenomenon has been reported for both adults and children (Robertson & Suci, 1980).

These findings can once again be readily explained within a framework where Agents initiate the construction of an event representation and/or evoke predictions about upcoming events. As demonstrated in Experiment 2, participants readily generated converging predictions about upcoming events based solely on the presentation of panels depicting an Agent in a preparatory state. Our analysis of completed actions in Experiment 4 showed that participants processed these panels faster when preceded by Agent First orders than Patient First orders. However, as we demonstrated in Experiment 5, this effect appeared to be motivated by the Agents—not serial order—since participants were faster to process a completed action when preceded by only an Agent in a preparatory state compared to only a Patient. Taken together, these results point to Agents being more central in the construction of event structure than Patients. This would suggest that independently of how event information is presented, through language, drawing, or film, internal representations of events are anchored to the Agents. Thus, the Agent becomes available sooner than the Patient when participants are asked to identify thematic roles.

This is perhaps most striking in the case of written passive constructions (Ferreira, 2003) where the Patient argument was presented significantly earlier than the Agent argument. In theory, this would have given participants more time to process the Patient than the Agent, which one might assume would make this semantic role easier to identify. Yet Ferreira (2003) found across several experiments that participants identified the Agent faster than the Patient, suggesting that Agents retained their primacy in event structure representations even in cases where the Patient was topicalized by the grammatical construction.

7.3. Effect of ordinal position and laying a foundation

Our findings across ordinal sequence position further inform about the relative processing of Agents and Patients, as well as their incorporation into a canonical serial order. First, viewing times

in both Experiments 3 and 4 were slowest at the start of sequences, and decreased across ordinal panel position as readers build up information about the narrative sequence (Cohn et al., 2012; Gernsbacher, 1985). Slower viewing times at the start of a sequence are a consequence of "laying a foundation" of information that the rest of the discourse then references (Gernsbacher, 1983, 1985, 1990). This function is further suggested since Agent panels, whether viewed first or second, steadily grew faster in viewing times across sequences in Experiment 4, suggesting that laying a foundation was less needed as the sequence progressed. In contrast, viewing times to Patients following Agents remained consistently fast across all panel positions, because their preceding Agents establish their role. This is unlike the Patients preceding Agents, which did not have their roles defined, and thus decrease as structure builds up over the sequence. Thus, semantic roles across sequence position further support that their relative processing indexes the access of event structure.

A comparable principle to "laying a foundation" may also influence the canonical serial order as an additional constraint to guide characters to precede actions in the Agent–Patient-Act orders found in various non-verbal expressions (Gershoff-Stowe & Goldin-Meadow, 2002; Goldin-Meadow et al., 2008). By providing all referential information first, the primary figures of an action can be established before binding them together in the action itself. Such a principle (e.g., Entities > Act) would conform to the general observation that discourse prefers to describe who is doing an action before describing the action itself (Primus, 1993; Sasse, 1987). Thus, again, this full canonical Agent–Patient-Act order results from the most efficient ways to detail events and their parts.

7.4. Prediction and facilitation

Finally, given that these experiments divided visually depicted events into their component parts, our results offer insights on the processing of real-life events, beyond their expression in communicative systems like language or manual expression. In Experiments 1 and 2 we found that Agents in preparatory actions were described as more active and more predictive about subsequent events than Patients. While Agent panels in Experiment 4 were predominantly viewed slower than Patients, we found that completed actions followed by Agent First orders were viewed faster than those preceded by Patient First orders. Finally, in Experiment 5, we found that panels depicting completed actions were viewed faster when followed by preparatory Agents than Patients. One explanation of these results is that Agents in preparatory actions facilitate the comprehension of subsequent events, while Patients do not.

Overall, these results suggest that semantic roles may be processed differently based on which part of an action they depict. We suggest that Agents require more processing in preparatory actions, which then facilitate comprehension at the event's completion. This preparatory information may help explain how people can anticipate subsequent events when they are presented as continuous films or real life (Zacks et al., 2007, 2011), as opposed to segmented panels. Here, cues from the Agents' preparatory actions inform the event's subsequent completion, as opposed to the relative passivity of Patients. While we did not investigate it in this study, we may hypothesize that, unlike Patients' passivity at an event's preparation, at its completion Patients would then require more processing, since they should undergo a change of state from the action (Kemmerer, 2012). Thus, after the completion of an event, we would expect Patients to be associated with greater processing costs compared to Agents. However, additional experimentation would be necessary to confirm this pattern.

7.5. Conclusions

In the current study we examined the proposal that Agents play a central role in the building of event structure. We found that participants interpret Agents as more active than Patients (Experiment 1) and lead participants to generate more predictions regarding upcoming events than Patients (Experiment 2). We found that Agents incur greater processing costs in preparatory states during on-line comprehension compared to Patients (Experiment 4). We also found that viewing times to Agents are modulated by their ordinal position within a narrative (Experiment 4). We demonstrated that participants were faster to process explicit depictions of events following the presentation of Agent–Patient orders compared to Patient-Agent orders (Experiment 4) and following the

presentation of single Agents compared to presentation of single Patients (Experiment 5). Altogether, our data strongly suggest that Agents play a unique role in generating predictions regarding upcoming events and may initiate the construction of event representations. These results contribute to explaining why previous studies have shown that participants preferentially attend to Agents over Patients under various experimental conditions.

Finally, we have argued that this finding may contribute to explaining the preference for placing Agents before Patients within human communication. Our results suggest that this pattern arises from the way in which semantic roles are incorporated into an event structure, guided by two primary constraints. First, the Agent–Patient order reflects an advantage for Agents over Patients as more integral to defining an event (i.e., an "Agent salience principle": Agent > Patient). Second, the fronting of referential information prior to describing an action reflects a preference for referential entities to appear prior to the events that bind them (i.e., "laying a foundation": Entities > Act). Together, these principles result in an Agent–Patient-Act order (i.e., [Agents > Patients] > Act). All told, this study demonstrates the usefulness of utilizing sequential visual narratives in exploring human cognition related to event processing.

Acknowledgments

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cogpsych.2013.07.002.

References

Aissen, J. (1999). Markedness and subject choice in optimality theory. *Natural Language and Linguistic Theory, 17*, 673–711.
Chatterjee, A. (2001). Language and space: Some interactions. *Trends in Cognitive Sciences, 5*(2), 55–61.
Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition, 2*(1), 79–116.
Chatterjee, A., Maher, L. M., & Heilman, K. M. (1995). Spatial characteristics of thematic role representation. *Neuropsychologia, 33*(5), 643–648.
Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events, and spatial representations. *Neuropsychologia, 37*(4), 395–402.
Clark, H. H. (1965). Some structural properties of simple active and passive clauses. *Journal of Verbal Learning and Verbal Behavior, 4*, 365–370.
Cohn, N. (2007). A visual lexicon. *Public Journal of Semiotics, 1*(1), 53–84.
Cohn, N. (2010). The limits of time and transitions: Challenges to theories of sequential image comprehension. *Studies in Comics, 1*(1), 127–147.
Cohn, N. (2013). Visual narrative structure. *Cognitive Science, 37*(3), 413–452.
Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2012). (Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension. *Cognitive Psychology, 65*(1), 1–38.
Dobel, C., Diesendruck, G., & Bölte, J. (2007). How writing system and age influence spatial representations of actions. *Psychological Science, 18*(6), 487–491.
Dryer, M. S. (2011). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Munich: Max Planck Digital Library.
Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*(2), 164–203.
Fincher-Kiefer, R. (1993). The role of predictive inferences in situation model construction. *Discourse Processes, 16*(1–2), 99–124.
Fincher-Kiefer, R. (1996). Encoding differences between bridging and predictive inferences. *Discourse Processes, 22*(3), 225–246.
Fincher-Kiefer, R., & D'Agostino, P. R. (2004). The role of visuospatial resources in generating predictive and bridging inferences. *Discourse Processes, 37*(3), 205–224.
Geminiani, G., Bisiach, E., Berti, A., & Rusconi, M. L. (1995). Analogical representation and language structure. *Neuropsychologia, 33*(11), 1565–1574.
Germeys, F., & d'Ydewalle, G. (2007). The psychology of film: Perceiving beyond the cut. *Psychological Research, 71*, 458–466.
Gernsbacher, M. A. (1983). *Memory for surface information in non-verbal stories: Parallels and insights to language processes.* Unpublished Dissertation, University of Texas at Austin, Austin, Texas.
Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology, 17*, 324–363.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Earlbaum.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language, 27*(6), 699–717.

Gernsbacher, M. A., Varner, K. R., & Faust, M. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 430–445.

Gershoff-Stowe, L., & Goldin-Meadow, S. (2002). Is there a natural order for expressing semantic relations? *Cognitive Psychology, 45*, 375–412.

Glanzer, M., Fischer, B., & Dorfman, D. (1984). Short-term storage in reading. *Journal of Verbal Learning and Verbal Behavior, 17*, 467–486.

Goldin-Meadow, S. (2003). *The resiliance of language: What gesture creation in deaf children can tell us about how all children learn language*. New York and Hove: Psychology Press.

Goldin-Meadow, S., & Feldman, H. (1977). The development of language-like communication without a language model. *Science, New Series, 197*(4301), 401–403.

Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Noneffect of parental input on language development. *Science, New Series, 221*(4608), 372–374.

Goldin-Meadow, S., So, W. C., Özyürek & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences, 105*(27), 9163–9168.

Greenberg, J. H. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of grammar* (pp. 73–113). Cambridge, MA: MIT Press.

Greeno, J. G., & Noreen, D. L. (1974). Time to read semantically related sentences. *Memory and Cognition, 2*, 117–120.

Haberlandt, K. (1984). Components of sentence and word reading times. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 219–251). Hillsdale, NJ: Erlbaum.

Hafri, A., Papafragou, A., & Trueswell, J. C. (2012). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General, 142*(3), 880–905.

Hawkins, J. A. (1983). *Word order universals*. New York: Academic Press.

Hoeks, J. C. J., Vonk, W., & Schriefers, H. (2002). Processing coordinated structures in context: The effect of topic-structure on ambiguity resolution. *Journal of Memory and Language, 46*(1), 99–119.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

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

Keenan, E. (1976). Towards a universal definition of "Subject". In C. Li (Ed.), *Subject and topic* (pp. 303–334). New York: Academic Press.

Kemmerer, D. (2012). The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's area. *Language and Linguistics Compass, 6*(1), 50–66.

Kieras, D. E. (1978). Good and bad structure in simple paragraphs: Effects on apparent theme, reading time, and recall. *Journal of Verbal Learning and Verbal Behavior, 17*, 13–28.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Science, 12*(2), 72–79.

Lasher, M. D. (1981). The cognitive representation of an event involving human motion. *Cognitive Psychology, 13*(3), 391–406.

Maass, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events. *Psychological Science, 14*(4), 296–301.

Magliano, J. P., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedent and causal consequence inferences. *Discourse Processes, 16*(1–2), 35–53.

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes, 22*, 199–224.

Magliano, J. P., Taylor, H., & Kim, H.-J. J. (2005). When goals collide: Monitoring the goals of multiple characters. *Memory and Cognition, 33*(8), 1357–1367.

Mandler, J. M., & Goodman, M. S. (1982). On the psychological validity of story structure. *Journal of Verbal Learning and Verbal Behavior, 21*, 507–523.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8*, 1–71.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(1), 82.

Michotte, A. (1946). *La Perception de la Causalité*. Leuven, Belgium: Publications Universitaires.

Miller, G. A., & McKean, K. O. (1964). A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology, 16*(4), 297–308.

Nakazawa, J. (2005). Development of manga (comic book) literacy in children. In D. W. Shwalb, J. Nakazawa, & B. J. Shwalb (Eds.), *Applied developmental psychology: Theory, practice, and research from Japan* (pp. 23–42). Greenwich, CT: Information Age Publishing.

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language, 27*(4), 399–415.

Primus, B. (1993). Word order and information structure: A performance based account of topic positions and focus positions. In J. Jacobs (Ed.), *Handbuch syntax* (pp. 880–895). Berlin and New York: de Gruyter.

Robertson, S. S., & Suci, G. J. (1980). Event perception by children in the early stages of language production. *Child Development, 51*(1), 89–96.

Sasse, H.-J. (1987). The thetic/categorical distinction revisited. *Linguistics, 25*, 511–580.

Segalowitz, N. S. (1982). The perception of semantic relations in pictures. *Memory and Cognition, 10*(4), 381–388.

Silverstein, M. (1976). Hierarchy of features and ergativity. In R. M. W. Dixon (Ed.), *Grammatical categories in Australian languages* (pp. 112–171). Canberra: Australian Institute of Aboriginal Studies.

Sitnikova, T., Holcomb, P. J., & Kuperberg, G. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience, 20*(11), 1–21.

Szewczyk, J. M., & Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language, 68*(4), 297–314.

Tan, E., & Diteweg, G. (1996). Suspense, predictive inference, and emotion in film viewing. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Eds.), *Suspense: Conceptualizations, theoretical analyses, and empirical explorations)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tannenbaum, P. H., & Williams, F. (1968). Generation of active and passive sentences as a function of subject or object focus. *Journal of Verbal Learning and Verbal Behavior, 7*(1), 246–250.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47*(1), 69–90.

van den Broek, P., Rapp, D., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes, 39*(2–3), 299–316.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory and Cognition, 19*, 95–112.

Verfaillie, K., & Daems, A. (1996). The priority of the agent in visual event perception: On the cognitive basis of grammatical agent–patient asymmetries. *Cognitive Linguistics, 7*(2), 131–148.

Weingartner, K. M., Guzmán, A. E., Levine, W. H., & Klin, C. M. (2003). When throwing a vase has multiple consequences: Minimal encoding of predictive inferencest. *Discourse Processes, 36*(2), 131–146.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., et al (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*(6), 651–655.

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of Cognitive Neuroscience, 23*(12), 4057–4066.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133*(2), 273–293.

Zubin, D. A. (1979). Discourse function of morphology: The focus system in German. In T. Givón (Ed.), *Syntax and semantics: Discourse and syntax* (pp. 469–504). New York: Academic Press.