TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science (2019) 1–19

ISSN:1756-8765 online
DOI: 10.1111/tops.12475

This article is part of the topic “Event-Predictive Cognition: From Sensorimotor via Conceptual to Language-Based Structures and Processes,” Martin V. Butz, David Bilkey, and Alistair Knott (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

From Event Representation to Linguistic Meaning

Ercenur Ünal,[a] Yue Ji,[b] Anna Papafragou[c]

[a]*Department of Psychology, Ozyegin University*
[b]*Department of Linguistics, University of Delaware*
[c]*Department of Linguistics, University of Pennsylvania*

Received 15 March 2019; received in revised form 25 September 2019; accepted 7 October 2019

Abstract

A fundamental aspect of human cognition is the ability to parse our constantly unfolding experience into meaningful representations of dynamic events and to communicate about these events with others. How do we communicate about events we have experienced? Influential theories of language production assume that the formulation and articulation of a linguistic message is preceded by preverbal apprehension that captures core aspects of the event. Yet the nature of these preverbal event representations and the way they are mapped onto language are currently not well understood. Here, we review recent evidence on the link between event conceptualization and language, focusing on two core aspects of event representation: event roles and event boundaries. Empirical evidence in both domains shows that the cognitive representation of events aligns with the way these aspects of events are encoded in language, providing support for the presence of deep homologies between linguistic and cognitive event structure.

Keywords: Event cognition; Event roles; Thematic roles; Event boundaries; Boundedness; Telicity

1. Introduction

A fundamental aspect of human cognition is the ability to perceive and organize the continuous stream of dynamic experience around us into discrete units known as *events*.

Correspondence should be sent to Anna Papafragou, Department of Linguistics, University of Pennsylvania, 3401-C Walnut St., Suite 300-C Wing, Philadelphia, PA 19104. E-mail: anna4@ling.upenn.edu

Segmenting events from continuous input is critical for interpreting and remembering our experiences, predicting others' actions, understanding event descriptions, and producing sentences or longer narratives about events in the world (Kurby & Zacks, 2008; Richmond & Zacks, 2017).

An event can be defined in multiple ways. For present purposes, we compile a definition from two sources that reflect two core aspects of an event representation. According to Elman (2009), an event is defined by the "set of participants, activities, and outcomes that are bound together by causal interrelatedness" (p. 572). In this view, the event participants and the relational structure that binds them are crucial for event representation. Zacks and Tversky (2001) further define an event as "a segment of time at a given location that is conceived by an observer to have a beginning and an end" (p. 3). This view highlights a spatiotemporal framework including duration, boundaries, and spatial structure as a core feature of an event. Event representations in memory, known as *event models*, capture information about these two core event aspects and store information about entities, objects, and the spatiotemporal framework of an event (Richmond & Zacks, 2017).

It has been proposed that, in order to segment ongoing activity into events, people form predictions about upcoming happenings based on a current event model (Event Segmentation Theory—Zacks, Speer, Swallow, Braver, & Reynolds, 2007; cf. also Radvansky & Zacks, 2014; Zwaan & Radvansky, 1998). When core features of the situation change, people can no longer predict upcoming happenings with precision and the current event model has to be updated into a new model. This change is experienced as an event boundary. The changes in the features that guide the detection of an event boundary can involve changes with respect to the event participants (e.g., cutting a cucumber into small pieces), changes along spatiotemporal characteristics (e.g., a person climbing up the stairs and entering into a building), or changes related to intentionality (e.g., deciding to watch a movie) or causality (e.g., breaking a vase). Importantly, scholars have proposed that sensitivity to these features comes from general cognitive biases that guide event apprehension (Shipley & Zacks, 2008; Zacks et al., 2007).

In addition to perceiving and experiencing events, people frequently communicate about the events they have perceived with others. How does the way people perceive events map onto the way they communicate about them? Influential theories of language production assume that speaking begins with a preverbal apprehension of broad details of an event that contains structured package of information, also known as a preverbal message (Levelt, 1989, cf. Bock, Irwin, & Davidson, 2004; Lashley, 1951; Paul, 1970; Wundt, 1970; see Papafragou & Grigoroglou, 2019, for a recent review). This preverbal message is constructed by drawing on the resources of the human mind that are responsible for conceptualizing different aspects of the world, including people, objects, places, time, and relations, among others. This preverbal message feeds information into the next stage of processing that takes into account various constraints on how entities, relations, and spatiotemporal information are encoded into language-specific lexical and syntactic structures in a way that ultimately gives rise to an utterance. Despite the fact that this view of language production has been highly influential in driving a large body of empirical research, several gaps remain in our understanding of both the early stages of preverbal event apprehension and how these preverbal event representations are mapped onto linguistic messages.

A major limitation of the current state of the art is that little is known about the nature of preverbal event representations—and to what extent they reflect the way different aspects of events are encoded in language (cf. also Papafragou, 2015). More specifically, several questions remain open with respect to how viewers extract information about core aspects of events, including event participants and spatiotemporal features, and how the information about these features contributes to the construction of structured event representations that are later mapped onto language. One possibility is that event representations in language and cognition are homologous such that the lexical and syntactic structures encoding different aspects of events build onto non-linguistic event representations (Jackendoff, 1983; Landau & Hoffman, 2005; Landau & Jackendoff, 1993). However, until recently, empirical evidence for this possibility has been limited due to the fact that the two systems have been studied separately, by different communities of scholars and using different methodologies. Without explicit evidence linking the mechanisms underlying the two systems, it is difficult to know whether there is indeed a link between linguistic and non-linguistic event representations, or whether other factors account for similarities between the two systems.

Here, we review empirical evidence on how events are represented in language and cognition to shed light onto some of these open questions on the nature of event representations and their links to language. We focus on two sub-domains, event roles and event boundaries, because they represent two core aspects of an event representation (Elman, 2009; Zacks & Tversky, 2001): Event roles capture relational information about the specific roles that event participants fulfill, and event boundaries encode critical spatial and temporal features of an event.

2. Event roles in language and cognition

As mentioned already, preverbal event apprehension has been assumed to represent entities or event participants as well as the causal relations among these entities. Linguistically, thematic roles describe how sentential constituents (i.e., noun or prepositional phrases) are related to a verb and capture relational information about who did what to whom (Fillmore, 1968; Jackendoff, 1990). Some of these thematic roles include the Agent, that is, the entity responsible for an action (e.g., *<u>A man</u> is eating*), the Patient, that is, the entity being affected by the action (e.g., *A man is eating <u>pizza</u>*), the Goal, that is, the end state of the action (e.g., *A man is putting pizza <u>into the oven</u>*), and the Instrument, that is, the means being used to create a change in the Patient (e.g., *A man is putting pizza into the oven <u>with a peel</u>*). Scholars have suggested that thematic roles in language map relatively directly onto elements of the underlying conceptual structure (Jackendoff, 1983, 1990).

Two sets of findings provide indirect support for the link between thematic roles and the underlying event concepts. First, a set of studies with adults has revealed that viewers can rapidly extract information about individual event components (Griffin & Bock, 2000) and can do so more rapidly from coherent scenes compared to incoherent scenes (Dobel, Gumnior, Bölte, & Zwitserlood, 2007). However, this line of work is limited by the fact that only a small number of event roles and event categories were investigated (cf. also Cohn & Paczynski, 2013; Dobel, Glanemann, Kreysa, Zwitserlood, & Eisenbeiss, 2010; Webb, Knott, & MacAskill, 2010; Zwitserlood, Bölte, Hofmann, Meier, & Dobel, 2018). A second piece of evidence comes from research with infants showing that infants demonstrate some understanding of conceptual correlates of thematic roles before they acquire language (for an overview see Göksun, Hirsh-Pasek, & Golinkoff, 2010; Wagner & Lakusta, 2009; cf. Cohen & Oakes, 1993; Golinkoff, 1975, 1981; Lakusta, Wagner, O'Hearn, & Landau, 2007; Woodward, 1998).

In a study that tested the representation of event roles more directly, we (Hafri, Papafragou, & Trueswell, 2013) investigated how viewers extract information about event roles after brief visual exposure to events. Participants briefly saw still images depicting two-participant events (e.g., a girl punching a boy) for either 37 or 73 ms and later responded with "yes" or "no" to verbal prompts about the Agent (e.g., "Is the girl performing the action?"), the Patient (e.g., "Is the boy being acted upon?"), the event (e.g., "Did you see punching?"), or a combination of the event and its participants (e.g., "The girl is punching the boy."). Results showed that participants could confirm a probe that was consistent with the event (e.g., when they responded to the "punching" prompt after seeing the punching event) and reject a probe that was inconsistent with the event (e.g., when they responded to the "punching" prompt after seeing a scaring event) under both durations. These findings suggest that viewers can quickly identify an event, its participants, and the combination of both even after a very brief visual exposure to an event.

A follow-up study tested whether information about Agent and Patient roles can be extracted while attention is occupied by an irrelevant task (Hafri, Trueswell, & Strickland, 2018). Viewers briefly saw the same still images as in the earlier study (Hafri et al., 2013) and identified which side of the screen a target (i.e., the one wearing a blue shirt or the one wearing a red shirt) was on. Importantly, in some trials, the target participant switched roles across trials (from Agent to Patient or from Patient to Agent), but in other trials the roles stayed the same. Viewers were faster to identify the side of the target participant when the role of the target stayed the same compared to when it switched, suggesting that Agent and Patient roles can be extracted even when role identification is not encouraged by task demands. Together, these findings suggest that viewers can rapidly and automatically extract event roles from simple two-participant events in a way that maps onto similar roles in language.

What about more complex events with multiple participants? In linguistic theory, thematic roles are ranked differently with respect to each other following a *Thematic Hierarchy*: Agents are considered most prominent, followed by Patients, then Goals, and Instruments are considered least prominent (even though this hierarchy has generated much discussion; Baker, 1997; Jackendoff, 1990; cf. Rissman & Majid, 2019; Rissman &

Rawlins, 2017; Rissman, Rawlins, & Landau, 2015). From a linguistic perspective, the Thematic Hierarchy can lead to the expectation that more prominent entities will be mentioned more frequently and in specific syntactic positions (Gernsbacher, 1989; Grosz et al., 1995; Meyer et al., 1998). From a conceptual perspective, it is reasonable to assume that more prominent entities are remembered better (Lakusta & Landau, 2005; Papafragou, 2010), detected faster (Griffin & Bock, 2000; Rensink et al., 1997), and named more easily (Dobel et al., 2007) despite variations in low-level perceptual properties such as size, shape, color, or luminance. A set of studies from our lab has investigated the link between linguistic and conceptual event roles focusing on complex events that involve multiple components. Using closely matched linguistic and non-linguistic measures of the salience of event components, we asked to what extent the relative prominence of event components in language and cognition conforms to the Thematic Hierarchy.

In one study, participants saw caused-motion events in which a person or an animal (Agent) made an object (Patient) move to a destination (Goal) using a tool or body part (Instrument) and had to describe the events (e.g., *An archer shooting an arrow towards a target with a bow*; Wilson, Ünal, Trueswell, & Papafragou, 2014). Participants mentioned Agents most frequently, followed by Patients, then Goals and Instruments, suggesting that the relative salience of event components in language follows the Thematic Hierarchy. Another group of participants in the same study completed a change-blindness task in which two versions of the same caused-motion events flickered with a gray mask between them. The only difference between the two versions of the event was the color of one event component. Participants were instructed to respond as soon as they detected the changing object. Participants were faster to detect changes to Patients (1,545 ms) compared to both Goals (1,877 ms) and Instruments (2,278 ms) and changes to Goals compared to Instruments. Changes to Agents were not detected the fastest (1,971 ms), presumably because the color change affected a property of the Agents (i.e., their clothes) as opposed to the Agents themselves. However, the relative salience of the remaining entities in this non-linguistic task also conformed to the Thematic Hierarchy.

In another study, participants viewed the same caused-motion events while their eye-movements were being recorded (Wilson, Papafragou, Bunger, & Trueswell, 2011). The participants' task was to move their eyes to a designated event component as soon as possible and to press a button once they did so. In the Agent condition, participants were instructed to look at the "the person or animal who was performing the action." In the Patient condition, participants were instructed to look at "the object directly affected by the action." In the Goal condition, participants were instructed to look at "the goal or destination of the action." Finally, in the Instrument condition, they were instructed to look at "the tool or body part used to make the action." In the Agent condition, looks to the Agents began diverging early on at around 120 ms from picture onset. Looks to the Patients in the Patient condition diverged slightly later at around 150 ms from picture onset. Looks to the Goals in the Goal condition and to the Instruments in the Instrument condition diverged much later (at 300 ms and 450 ms from picture onset, respectively). One might ask whether the differences in the speed of identification of event components could be explained by lower-level perceptual features of the stimuli such as the size of a component. A follow-up analysis tested this possibility by assessing whether time taken to fixate on a component correlated with the size of the component (as measured as a percentage of image area using the Tobii Studio AOI tool). No such correlations were found. Together, these findings are consistent with the conclusion that the relative salience of event components varies and provide more direct evidence for a link between thematic roles in language and event conceptualization.

In a developmental extension of these studies, we asked whether the asymmetries in event role prominence also characterize the period during which children acquire language (Ünal, Trueswell, & Papafragou, 2017). Furthermore, since the studies reviewed so far were all conducted with speakers of English, we tested young learners of both English and Turkish to probe whether asymmetries in event role prominence generalize to other languages that vary in the realization of thematic roles within syntactic positions. In English, Agents typically appear as subject noun phrases, Patients as direct object noun phrases, and Goals or Instruments as adpositional phrases (see example in 1 below). There is more surface variation in Turkish in the encoding of Agents, Goals, and Instruments. Agents can be encoded as subject noun phrases (see 2a) or can be dropped (see 2b) because the verb is marked for person and number. Goals can be encoded either as postpositional phrases (see 2b) or noun phrases marked with dative case (see 2a). Instruments can be encoded as noun phrases in the commitative case (see 2b) or verbs (see 2a).

(1)	A woman hit the ball into a basket with a tennis racquet

(2a)	Adam$_{\text{[AGENT]}}$	sepete$_{\text{[GOAL]}}$	süpürüyor$_{\text{[INSTRUMENT]}}$	yaprakları$_{\text{[PATIENT]}}$
	The man	basket-DAT	sweep-PRES	leaves-ACC

(2b)	Topu$_{\text{[PATIENT]}}$	çantanın içine$_{\text{[GOAL]}}$	süpürgeyle$_{\text{[INSTRUMENT]}}$	atıyor
	ball-ACC	inside (the) bag	sweeper-COM	throw-PRES-3SG

English- and Turkish-speaking 3- and 4-year-old children completed the same linguistic description and change blindness tasks as in the earlier paradigm (Wilson et al., 2014). In the linguistic task, Turkish-speaking children mentioned Agents less frequently (44%) than English-speaking children (83%). Patients, Goals, and Instruments were mentioned equally frequently in both languages, with Instruments (Turkish 24%, English 20%) being mentioned less than either Patients (Turkish 78%, English 83%) or Goals (Turkish 59%, English 61%). In the change blindness task, both English- and Turkish-speaking children were more accurate in detecting changes to Patients (Turkish 63%, English 73%) and Goals (Turkish 59%, English 77%) compared to Instruments (Turkish 36%, English 45%). Both language groups had comparable levels of accuracy for Agent changes (Turkish 54%, English 55%; as in the adult data, these levels were somewhat low because of the insalience of a color change in the Agent's clothing). These findings suggest that Patients and Goals tend to be similarly prioritized over Instruments in both language and cognition. Furthermore, the relative salience of event components was similar across young learners of English and Turkish, despite the cross-linguistic differences in the surface encoding of these entities.

An important aspect of these studies that needs to be taken into account is the fact that they involve a comparison between roles that are syntactically encoded as arguments (Agents and Patients) to roles that are often encoded as adjuncts (Goals and Instruments). Thus, one might ask whether the asymmetries in event roles reported above might be driven by the surface syntactic encoding of thematic roles as opposed to thematic relations per se. Three pieces of evidence suggest that this alternative explanation is unlikely. First, asymmetries observed in English were also observed in Turkish, which shows more flexibility in both word order (Erguvanlı, 1984) and, critically, the syntactic positions in which event roles are encoded (Göksel & Kerslake, 2005). Second, even though these event roles differed in terms of their syntactic encoding, some of the role prominence asymmetries observed in the present studies were independent of these syntactic variations. For example, Agents and Patients are both encoded as arguments, but their relative salience varied. On the other hand, Patients are encoded as arguments and Goals as adjuncts, but Patients were not always more salient than Goals. Finally, there is independent evidence from the sentence processing literature for a dissociation of syntactic and semantic components of thematic roles (Bornkessel, Schlesewsky, & Friederici, 2003; Scheepers, Hemforth, & Konieczny, 2000).

Summarizing, the studies reviewed above demonstrate that viewers rapidly identify the event type (Dobel et al., 2007; Zwitserlood et al., 2018) and event roles relevant to language for simple two-participant events (Griffin & Bock, 2000; Hafri et al., 2013). Furthermore, the roles of event participants are identified somewhat automatically (Hafri et al., 2018). Recent work extends these findings to complex multi-participant (Wilson et al., 2011) events and other event categories such as caused-motion events. Furthermore, event components are similarly prioritized in language and cognition (Ünal et al., 2017; Wilson et al., 2014). Taken together, these findings suggest that preverbal event apprehension represents information about event participants and their causal relatedness in ways that map onto the way the roles of these event participants are encoded in language.

3. Event boundaries in language and cognition

A second core kind of information represented in event models is the information concerning spatial and temporal features of the event, including event boundaries. Event boundaries are detected when a crucial set of features of the situation change. The changes that suggest an event boundary can be perceptual, such as a new spatial location (e.g., moving from a space shuttle to an aircraft, Magliano, Miller, & Zwaan, 2001), a new direction or speed of motion (Zacks, 2004), new body positions (Newtson, Engquist,

& Bois, 1977), or even the appearance of new event participants (Zacks, Speer, & Reynolds, 2009). The changes can also be conceptual and may relate to how people use their knowledge of goals, causes, and effects, and so on (e.g., a researcher arriving at a conclusion; Ratcliff & Lassiter, 2007; Zacks & Tversky, 2001).

Event boundaries have a privileged status in event cognition. For instance, objects and actions at an event boundary are retrieved more easily than those present at any moment between event boundaries (Kurby & Zacks, 2008; Newtson & Engquist, 1976; Swallow, Zacks, & Abrams, 2009), and introducing a visual distractor to an event boundary causes more disruption to event recall than introducing a distractor between event boundaries (Boltz, 1992). Furthermore, event boundaries serve as anchors for comprehending, learning, and describing events (e.g., Boltz, 1995; see Radvansky & Zacks, 2017 for a review). Event endpoints, in particular, have been argued to be critical for how events are represented. When participants view motion events in which one entity moves from one reference point (source) to another (goal), for instance, a butterfly flying from a shed to a lamppost, they tend to remember goals better than sources (Lakusta & Landau, 2012; Papafragou, 2010). When people compare two events, the resultant state (e.g., whether a ball knocked over a tower or just a few blocks from the tower structure) tends to have more psychological weight than other event features (e.g., the moving direction of the ball) (He & Arunachalam, 2016). When observers track and estimate the number of overlapping motions, asynchronous endings but not asynchronous beginnings have an effect on estimation (Ongchoco & Scholl, 2019). The resultant state is especially important for infants' processing and interpretation of others' actions. Infants before their first birthday react differently to expected and unexpected outcomes (Bíró, Verschoor, Coalter, & Leslie, 2014; Csibra, Bíró, Koós, & Gergely, 2003; Jovanovic, et al., 2007). In one study, after a brief exposure to unfamiliar actions (e.g., the movement of a mechanical claw) that led to a clear resultant state (e.g., a toy being lifted up by the claw), infants were able to anticipate the outcome despite the novelty of the inanimate agents; however, infants failed to form predictions if there was no clear result from the actions (e.g., the toy remained still after the claw's grasping action; Adam & Elsner, 2018; see also Csibra, 2008).

Evidence for the salience of event endpoints is also found in how people describe events. Many studies cross-linguistically have revealed that children and adults are less likely to mention motion sources (*The butterfly flew from the shed* . . .) compared to endpoints or goals (. . . *to the lamppost*) when describing motion events (e.g., see Johanson, Selimis, & Papafragou, 2019; Lakusta & Landau, 2005; Regier & Zheng, 2007). This bias to encode endpoints can be constrained by tense or aspect. The choice of tense and aspect may put constraints on the attentional bias for endpoints. When comprehending a perfective sentence (e.g., *John had opened the bottle*), people focus on the resultant state (e.g., an opened bottle) but when comprehending an imperfective sentence (e.g., *John was opening the bottle*), people show no preference between an ongoing and a completed state (Madden & Zwaan, 2003; Madden-Lombardi, Dominey, & Ventre-Dominey, 2017; cf. Kang, 2015; Krass & Altmann, 2017).

Despite the extensive literature on event segmentation and the salience of endpoints, the *nature* of event endpoints has received limited attention. Some events have a natural, inherent endpoint and are temporally *bounded* (e.g., crack an egg into a bowl, eat a sandwich), while others are unspecified about when they come to an end and are temporally *unbounded* (e.g., stir an egg in a bowl, eat cheerios). This distinction is systematically encoded in language, mainly through verbs that denote different types of actions (e.g., *crack* vs. *stir*) and the verb's argument that denotes the nature of the affected object(s) (e.g., *eat a sandwich* vs. *eat cheerios*) (Filip, 2004, 2012; Rothstein, 2004; van Hout, de Swart, & Verkuyl, 2005). Bounded events are considered as developments leading to a "built-in terminal point" (Comrie, 1976), "climax" (Vendler, 1957), or "culmination" (Parsons, 1990). By contrast, unbounded events have a homogenous internal structure (Hinrichs, 1985; Krifka, 1989, 1998; Taylor, 1977) and may terminate at any arbitrary moment. The linguistic dichotomy between bounded and unbounded events at some level can be a candidate for a semantic universal (von Fintel & Matthewson, 2008), even though it can be expressed in different ways across languages (Bar-el, Davis, & Matthewson, 2005; Botne, 2003). This perspective strongly suggests the possibility that boundedness is grounded in the conceptual representation of events (Goldberg, 1995; Jackendoff, 1990; Pustejovsky, 1991).

Research on sign languages has provided some support for this possibility. Some sign languages tend to express boundedness in an iconic way. That is, signs denoting bounded events (e.g., decide) have a salient visual boundary and tend to have abrupt, single-stroke movements, whereas signs for unbounded events (e.g., think) do not have a salient visual boundary and tend to consist of repetitive movements (Malaia, Wilbur, & Milcovič, 2013; Wilbur, 2003). Recent work shows that the iconic mapping between the meaning (i.e., boundedness) and form (visual boundary) is also accessible to hearing non-signers (Strickland et al., 2015). Furthermore, neuroimaging studies have revealed the connection between the perception of event boundaries and their expression in language by showing that overlapping brain areas are activated when people use perceptual cues for event segmentation and process signs denoting bounded or unbounded events (e.g., Malaia, Renaweera, Wilbur, & Talavage, 2012; see Malaia, 2014 for a review). Nevertheless, both of these findings rely on the interpretation of communicated symbols (signs), not the representation of boundedness in event cognition per se.

Recent studies from our laboratory provide the most direct piece of evidence for the cognitive basis of the linguistic notion of boundedness by comparing bounded and unbounded events in non-linguistic tasks. One study (Ji & Papafragou, 2017a) adopted a category identification task. In the training phase, participants watched videos of paired, minimally different bounded and unbounded events (e.g., a girl dressed a teddy bear vs. a girl patted a teddy bear), with videos of one event category surrounded by a red frame (indicating that they belonged to the target category). At test, participants watched videos of new events that they did not see before (e.g., a girl stirred yogurt) and had to decide whether these videos could get a red frame or not. The results showed that viewers were able to form categories corresponding to bounded and unbounded events. A follow-up study showed that the categorization process truly tracked whether the events had inherent endpoints, not whether the endpoints had actually been reached in the videos (i.e., whether the event was finished): Even when presented with truncated videos that included only the beginning and midpoint of the events, participants could draw the bounded vs. unbounded distinction. Further manipulations eliminated the possibility that the categorization results relied on describing the events: Even when participants' access to linguistic encoding was suppressed by a secondary counting task throughout the training phase, they succeeded in identifying both bounded and unbounded event categories. Furthermore, an additional study established that bounded and unbounded stimuli were treated as equally intentional so whatever differences emerged in categorization were not due to differences in intentionality. Together, these findings suggest that viewers are sensitive to the property of boundedness, such that event cognition tracks abstract dimensions of event structure.

In a developmental extension of this study with 4-to-5-year-old children, we modified the boundedness task by replacing the red frame with a star marking the target category and presenting pairs of bounded and unbounded events side by side on the screen during the training phase (Ji & Papafragou, 2017b). At test, children had to decide whether videos of new events could get a star. The results showed that children could identify the category of bounded events as they assigned a star to new bounded events and not to new unbounded events. However, unlike adults, children failed to form the unbounded event category—their assignment of the star was random in the testing phase. Further tests confirmed that tracking event boundedness differed from tracking event completion (children, just like adults, could generalize the bounded event category to truncated videos where the ending of the event was not visible). In sum, children from an early age know that some events have inherent endpoints and treat event boundedness as distinct from event completion.

An interesting question is why children failed to extract the class of unbounded events. Notice that, even for adults, the category of bounded events was identified with greater ease compared to that of unbounded events across all of our studies. It is possible that defined endpoints carry richer information, attract more attention, and thus make bounded events easier to individuate, track, compare to each other, and generalize over. Support for this possibility comes from a recent eye-tracking study by Sakarias and Flecken (2018) showing that the endpoints of bounded events are more "attention grabbing" compared to those of unbounded events. We are currently pursuing this topic in ongoing work.

To the extent that adult viewers are sensitive to the boundedness distinction, it should be possible to detect consequences for the way internal event structure is processed across event types. To test this possibility, a further study (Ji & Papafragou, 2018) introduced brief interruptions to block the temporal midpoints or endpoints of bounded and unbounded events, and examined how viewers reacted to such mid-interruptions and end-interruptions of event flow. The study adopted a variant of the "picky puppet task" (Waxman & Gelman, 1986). Participants were presented with a picky girl who liked videos containing either mid-interruptions or end-interruptions. The task was to identify which videos the girl would like within a new set. The results showed that participants who watched videos of bounded events performed better when the picky girl liked mid-interruptions compared to end-interruptions, but participants exposed to videos of unbounded events performed equally well in identifying preferences for either type of interruption. This suggests that, for bounded events, blocking the endpoint was more disturbing for viewers (and hence less acceptable as a "preference") compared to blocking the midpoint; however, for unbounded events, interruptions at the two time points were treated largely identically. These findings provide new evidence for the importance of event endpoints, extending previous literature (Lakusta & Landau, 2005; Papafragou, 2010; Regier & Zheng, 2007). More important, these results demonstrate that the salience of endpoints depends on event boundedness: Endpoints are weighed more heavily than midpoints only in bounded events (that have a finely differentiated internal structure) but not in unbounded events (that have a homogeneous structure).

Summarizing, these findings suggest that event cognition is sensitive to the linguistic concept of boundedness, which captures the nature of event endpoints as well as the internal temporal structure of events. These studies provide empirical evidence for the broadly accepted—but rarely tested—assumption that "[t]he notions 'bounded' and 'unbounded' belong to a finite set of primitives that characterizes parts of conceptual structure" (Filip, 1993, p. 10; cf. also Folli & Harley, 2006). The present findings are consistent with the idea that temporal organization of events play a crucial role in event representation in both language and cognition (cf. McRae, Brown, & Elman, in press).

4. Conclusions

In this article, we have reviewed a growing body of empirical evidence addressing the relation between event representations in language and cognition, focusing on two core aspects of events: event roles and event boundaries. Our goal was to assess how these aspects of events are represented non-linguistically, and how non-linguistic event representations map onto linguistic encoding, as a way of shedding light onto the open question of how the linguistic and cognitive systems interface with each other.

A key conclusion from the present review is that the cognitive representation of events is sensitive to aspects of events necessary for linguistic encoding. Studies on the representation of event roles in language and cognition have shown that event components corresponding to linguistic thematic roles can be identified rapidly in both simple events with two-participants (Hafri et al, 2013, 2018) and in more complex events involving additional participants (Wilson et al., 2011, 2014). Furthermore, studies with both adults and children have revealed that viewers can use the linguistic notion of boundedness as a relevant dimension for forming event categories (Ji & Papafragou, 2017a, 2017b). The fact that cognitive representations of both roles and event boundaries are sensitive to these linguistic distinctions provides a first piece of evidence for the presence of a homology between event representations in language and cognition. The present body of evidence is also consistent with the view that event models activated in working memory while perceiving and experiencing events and linguistic representation of events share similar structures (Knott & Takac, in press).

A further key conclusion is that the relative salience of certain aspects of events varies in similar ways across language and cognition. For example, viewers are faster or more accurate in identifying those event roles that are more likely to be mentioned in linguistic production (Ünal et al., 2017; Wilson et al., 2011, 2014). Furthermore, event endpoints tend to be encoded more frequently in language production (Lakusta & Landau, 2005; Papafragou, 2010), remembered better (Lakusta & Landau, 2012; Papafragou, 2010), and given greater weight when forming structured event representations compared to other temporal slices of events (Ji & Papafragou, 2018). The presence of similar asymmetries across language and cognition provides a second piece of evidence supporting the idea that the linguistic encoding of different aspects of events reflects (and possibly builds on) the organization of non-linguistic event representations. Converging evidence for this argument comes from work with preverbal infants demonstrating that their event representations are organized in ways reflecting linguistic categories such as Sources, Goals and other related roles (Cohen & Oakes, 1993; Golinkoff, 1975, 1981; Lakusta et al., 2007; Wagner & Lakusta, 2009; Woodward, 1998; Yin & Csibra, 2015), and notions such as boundedness (He & Arunachalam, 2016). The presence of these homologies between language and cognition has important implications for the origins of linguistic structures and the continuity of non-verbal event representations throughout development. Our understanding of these issues will benefit greatly from future developmental work jointly studying event language and cognition.

At present, several questions remain open for further research on events and language. First, current empirical evidence on the relative salience of event roles and event boundaries comes from the domain of change of state events and motion events. It remains an open question whether the asymmetries reported in current work generalize to other classes of events and other event participants. For instance, the source-goal asymmetry reported in prior work on motion memory (Lakusta & Landau, 2005; Papafragou, 2010) is only found when a motion event has an animate—but not an inanimate—agent (Lakusta & Landau, 2012). Thus, it would be interesting to investigate how asymmetries in the salience of aspects of events change across different types of events that involve different types and numbers of event components.

A second important issue is that people speak about events in communicative contexts. An intriguing open question is how the relation between event representations in language and cognition makes contact with pragmatics, such as features of the context or goals of the speaker. Recent work suggests that frequency of mention of certain event components, such as instruments (Grigoroglou & Papafragou, 2019) or sources of motion (Do, Papafragou, & Trueswell, 2019), change depending on the context. Thus, it would be interesting to test whether and how pragmatic constraints interact with the prominence of event components. Relatedly, when describing events in communicative contexts, speakers also often use non-verbal means such as gesture to encode aspects of events (ter Bekke, Özyürek, & Ünal, 2019; Kita & Özyürek, 2003). It would be important to assess whether the asymmetries reported in current work also emerge when events are encoded in other modalities.

Third, most of the research reported here has investigated the link between event representations in language and cognition with speakers of English. Nevertheless, as alluded to already, languages vary considerably in how different aspects of events are encoded, including event roles and boundedness. This cross-linguistic variation raises issues for the assumption that sensitivity to the features that guide event apprehension serve as general cognitive biases (e.g., Boroditsky, 2001; Levinson, 2003). Our own work on the relative salience of event roles in English- and Turkish-speaking children shows that children extract information about event roles similarly despite cross-linguistic differences in the encoding of event roles (Ünal et al., 2017). Although these findings cohere with a broader perspective on the nature of the interactions between language and cognition (Gleitman & Papafragou, 2016; Ünal & Papafragou, 2016), further cross-linguistic research with both infants, children, and adults in the domain of events and within event subdomains is necessary to evaluate this possibility more directly.

Finally, from the present description, it might appear as though cognitive and linguistic representations of events capture observable occurrences in the physical world in a fairly direct way. However, we know that the way we represent events combines input from multiple processes involved in cognition and at multiple levels of abstraction. Broadly speaking, this involves perceptual input (e.g., a change in object states) that is mainly processed in a bottom-up fashion, as well as higher-order information such as one's knowledge about goals, causes and effects, and so on that affects event representation in a top-down fashion (Zacks & Tversky, 2001). An intriguing question is how processes operating over distinct types of event information work together (e.g., Hard, Tversky, & Lang, 2006; Levine, Hirsh-Pasek, Pace, & Golinkoff, 2017; Ünal & Papafragou, 2019; Zacks, 2004). It is possible that higher level knowledge can affect the way people select and allocate attention to various perceptual features (Zacks, 2004; see also Baird & Baldwin, 2003). In our own ongoing work, we have found that the way people place event boundaries depends partly on how people interpret the goals of the agent within an event (A. Mathis & A. Papafragou, unpublished data). Future research needs to investigate how perceptual information about object states and conceptual understanding of goals and intentions are integrated into dynamically unfolding event representations.

Acknowledgments

We acknowledge support from a University of Delaware Doctoral Fellowship (Y.J.) and NSF grant 1632849 (A.P.).

References

Adam, M., & Elsner, B. (2018). Action effects foster 11-month-olds' prediction of action goals for a non-human agent. *Infant Behavior and Development*, *53*, 49–55. https://doi.org/10.1016/j.infbeh.2018.09.002

Baird, J. A., & Baldwin, D. (2003). Making sense of human behavior: Action parsing and intentional inference. In B. Malle, L. J. Moses, & D. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 193–206). Cambridge, MA: MIT Press.

Baker, M. C. (1997). Thematic roles and syntactic structure. In L. Haegeman (Ed.), *Handbook of generative syntax* (pp. 73–137). Dordrecht, the Netherlands: Kluwer.

Bar-El, L., Davis, H., & Matthewson, L. (2005). On non-culminating accomplishments. In L. Bateman & C. Ussery (Eds.), *Proceedings of the North Eastern Linguistics Society* (vol. *35*, pp. 87–102). Amherst, MA: GLSA.

Bíró, S., Verschoor, S., Coalter, E., & Leslie, A. M. (2014). Outcome producing potential influences twelve-month-olds' interpretation of a novel action as goal-directed. *Infant Behavior and Development*, *37*, 729–738. https://doi.org/10.1016/j.infbeh.2014.09.004

Bock, K., Irwin, D., & Davidson, D. (2004). Putting first things first. In J. Henderson & F. Ferreira (Eds.), *The interface between language, vision and action: Eye movements and the visual world* (pp. 249–317). New York: Psychology Press.

Boltz, M. G. (1992). Temporal accent structure and the remembering of filmed narratives. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 90–105. https://doi.org/10.1037/0096-1523.18.1.90

Boltz, M. G. (1995). Effects of event structure on retrospective duration judgments. *Perception and Psychophysics*, *57*, 1080–1096. https://doi.org/10.3758/BF03205466

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2003). Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing. *Language and Cognitive Processes*, *18*, 269–298. https://doi.org/10.1080/01690960244000018

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conception of time. *Cognitive Psychology*, *43*, 1–22. https://doi.org/10.1006/cogp.2001.0748

Botne, R. (2003). To die across languages: Towards a typology of achievement verbs. *Linguistic Typology*, *7*, 233–278. https://doi.org/10.1515/lity.2003.016

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology*, *29*, 421–433. https://doi.org/10.1037//0012-1649.29.3.421

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, *67*(3), 73–97. https://doi.org/10.1016/j.cogpsych.2013.07.002

Comrie, B. (1976). *Aspect: An introduction to the study of verb aspect and related problems*. Cambridge, UK: Cambridge University Press.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, *107*, 705–717. https://doi.org/10.1016/j.cognition.2007.08.001

Csibra, G., Bíró, B., Koós, S., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*, 111–133. https://doi.org/10.1207/s15516709cog2701_4

Do, M., Papafragou, A., & Trueswell, J. (2019). The goal bias in memory and language: Explaining the asymmetry. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 268–274). Montreal: Cognitive Science Society.

Dobel, C., Glanemann, R., Kreysa, H., Zwitserlood, P., & Eisenbeiss, S. (2010). Visual encoding of coherent and non-coherent scenes. In E. Pedersen & J. Bohnemeyer (Eds.), *Event representation in language: Encoding events at the language cognition interface* (pp. 189–215). Cambridge, UK: Cambridge University Press.

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

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33*(4), 547–582. https://doi.org/10.1111/j.1551-6709.2009.01023.x

Erguvanlı, E. E. (1984). *The function of word order in Turkish grammar*. Berkeley: University of California Press.

Filip, H. (1993). Aspect, situation types and nominal reference. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.

Filip, H. (2004). The telicity parameter revisited. *Semantics and Linguistic Theory*, *14*, 92–109. https://doi.org/10.3765/salt.v0i0.2909

Filip, H. (2012). Lexical aspect. In R. I. Binnich (Ed.), *The Oxford handbook of tense and aspect* (pp. 721–751). Oxford, UK: Oxford University Press.

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

Folli, R., & Harley, H. (2006). On the licensing of causatives of directed motion: Waltzing Matilda all over. *Studia Linguistica*, *60*, 121–155. https://doi.org/10.1111/j.1467-9582.2006.00135.x

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, *32*, 99–156. https://doi.org/10.1016/0010-0277(89)90001-2

Gleitman, L. R., & Papafragou, A. (2016). New perspectives on language and thought. In K. Holyoak, & R. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (2nd ed) (pp. 543–568). New York: Oxford University Press.

Göksel, A., & Kerslake, C. (2005). *Turkish: An essential grammar*. New York: Routlage.

Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2010). Trading Spaces: Carving up events for learning language. *Perspectives on Psychological Science*, *5*, 33–42. https://doi.org/10.1177/1745691609356783

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Golinkoff, R. M. (1975). Semantic development in infants: The concepts of agent and recipient. *Merrill-Palmer Quarterly*, *21*, 181–193. Retrieved from https://www.jstor.org/stable/23084058.

Golinkoff, R. M. (1981). The case for semantic relations: Evidence from the verbal and nonverbal domains. *Journal of Child Language*, *78*, 413–438. https://doi.org/10.1017/S0305000900003275

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

Grigoroglou, M., & Papafragou, A. (2019). Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology*, *55*, 951–966. https://doi.org/10.1037/dev0000693

Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, *21*, 203–225.

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

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

Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition*, *34*, 1221–1235. https://doi.org/10.3758/BF03193267

He, X., & Arunachalam, S. (2016). How event endstates are conceptualized. Poster presented at the 29th Annual CUNY Conference on Human Sentence Processing, Gainesville, FL.

Hinrichs, E. (1985). A compositional semantics for Aktionsarten and NP reference in English. Unpublished doctoral dissertation, The Ohio State University.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

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

Ji, Y., & Papafragou, A. (2017a). Viewers' sensitivity to abstract event structure. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 594–599). Austin, TX: Cognitive Science Society.

Ji, Y., & Papafragou, A. (2017b). Children's sensitivity to abstract event structure. Poster presented at the 42nd Annual Boston University Conference on Language Development, Boston, MA.

Ji, Y., & Papafragou, A. (2018). Midpoints and endpoints in event perception. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 1874–1879). Austin, TX: Cognitive Science Society.

Johanson, M., Selimis, S., & Papafragou, A. (2019). The Source-Goal asymmetry in spatial language: Language-general vs. language-specific aspects. *Language, Cognition and Neuroscience*, *34*, 826–840. https://doi.org/10.1080/23273798.2019.1584323

Jovanovic, B., Király, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (2007). The role of effects for infants' perception of action goals. *Psychologia: An International Journal of Psychology in the Orient*, *50*, 273–290. https://doi.org/10.2117/psysoc.2007.273

Kang, X. (2015). Establishing object-state representation in language comprehension: Evidence from picture verification, eye-tracking and ERPs. Unpublished doctoral dissertation, University of York, UK.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32. https://doi.org/10.1016/S0749-596X(02)00505-3

Knott, A., & Takac, M. (in press). Roles for event representations in sensorimotor experience, memory formation and language processing. *Topics in Cognitive Science*.

Krass, K., & Altmann, G. T. M. (2017). Affordances and end states: How verb type and tense affect event comprehension. Talk presented at the 30th Annual CUNY Conference on Human Sentence Processing, Boston, MA.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expression* (pp. 75–115). Dordrecht, the Netherlands: Foris Publications.

Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and grammar* (pp. 197–235). Dordrecht, the Netherlands: Kluwer.

Kurby, C. A., & Zacks, J. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*, 72–79. https://doi.org/10.1016/j.tics.2007.11.004

Lakusta, L., & Landau, B. (2005). Starting at the end: the importance of goals in spatial language. *Cognition*, *96*, 1–33. https://doi.org/10.1016/j.cognition.2004.03.009

Lakusta, L., & Landau, L. (2012). Language and memory for motion events: Origins of the asymmetry between goal and source paths. *Cognitive Science*, *36*, 517–544. https://doi.org/10.1111/j.1551-6709.2011.01220.x

Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development*, *3*, 179–197. https://doi.org/10.1080/15475440701360168

Landau, B., & Hoffman, J. E. (2005). Parallels between spatial cognition and spatial language: Evidence from Williams syndrome. *Journal of Memory and Language*, *53*(2), 163–185. https://doi.org/10.1016/j.jml.2004.05.007

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16*, 255–265. https://doi.org/10.1017/S0140525X00029733

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

Levelt, W. (1989). *Speaking*. Cambridge, MA: MIT Press.

Levine, D., Hirsh-Pasek, K., Pace, A., & Golinkoff, R. (2017). A goal-bias in action: The boundaries adults perceive in events align with sites of actor intent. *Journal of Experimental Psychology*, *43*, 916–927. https://doi.org/10.1037/xlm0000364

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, UK: Cambridge University Press.

Madden, C., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & Cognition*, *31*, 663–672. https://doi.org/10.3758/BF03196106

Madden-Lombardi, C., Dominey, P. F., & Ventre-Dominey, J. (2017). Grammatical verb aspect and event roles in sentence processing. *PLoS ONE*, *12*, 1–13. https://doi.org/10.1371/journal.pone.0189919

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology*, *15*, 533–545. https://doi.org/10.1002/acp.724

Malaia, E. (2014). It still isn't over: Event boundaries in language and perception. *Language and Linguistics Compass*, *8*, 89–98. https://doi.org/10.1111/lnc3.12071

Malaia, E., Renaweera, R., Wilbur, R., & Talavage, T. (2012). Event segmentation in a visual language: Neural bases of processing American Sign Language predicates. *NeuroImage*, *19*, 4094–4101. https://doi.org/10.1016/j.neuroimage.2011.10.034

Malaia, E., Wilbur, R., & Milcovič, M. (2013). Kinematic parameters of signed verbs. *Journal of Speech Language and Hearing Research*, *56*, 1677–1688. https://doi.org/10.1044/1092-4388(2013/12-0257)

McRae, K., Brown, K. S., & Elman, J. L. (in press). Prediction-based learning and processing of event knowledge. *Topics in Cognitive Science*.

Meyer, A. S., Sleiderink, A., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*, B25–B33. https://doi.org/10.1016/S0010-0277(98)00009-2

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12*, 436–450. https://doi.org/10.1016/0022-1031(76)90076-7

Newtson, D., Engquist, G. A., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology*, *35*, 847–862. https://doi.org/10.1037/0022-3514.35.12.847

Ongchoco, J. D. K., & Scholl, B. J. (2019). Did that just happen? Event segmentation influences enumeration and working memory for simple overlapping visual events. *Cognition*, *187*, 188–197. https://doi.org/10.1016/j.cognition.2019.01.002

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, *34*, 1064–1092. https://doi.org/10.1111/j.1551-6709.2010.01107.x

Papafragou, A. (2015). The representation of events in language and cognition. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 327–345). Cambridge, MA: MIT Press.

Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: Evidence from event encoding. *Language, Cognition and Neuroscience*, *34*(9), 1117–1128. https://doi.org/10.1080/23273798.2019.1589540

Parsons, T. (1990). *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.

Paul, H. (1970). The sentence as the expression of the combination of several ideas. In A. L. Numenthal (trans.), *Language and psychology: Historical aspects of psycholinguistics* (pp. 20–31). New York: Wiley.

Pustejovsky, J. (1991). The syntax of event structure. *Cognition*, *41*, 47–81. https://doi.org/10.1016/0010-0277(91)90032-Y

Radvansky, G. A., & Zacks, J. M. (2014). *Event cognition*. New York: Oxford University Press.

Radvansky, G., & Zacks, J. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, *17*, 133–140. https://doi.org/10.1016/j.cobeha.2017.08.006

Ratcliff, J. J., & Lassiter, G. D. (2007). On the induction and consequences of variation in behavior perception. *Current Psychology*, *26*, 16–36. https://doi.org/10.1007/s12144-007-9003-9

Regier, T., & Zheng, M. (2007). Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science*, *31*, 705–719. https://doi.org/10.1080/15326900701399954

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

Richmond, L. L., & Zacks, J. M. (2017). Constructing experience: Event models from perception to action. *Trends in Cognitive Sciences*, *21*, 962–980. https://doi.org/10.1016/j.tics.2017.08.005

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

Rissman, L., & Rawlins, K. (2017). Ingredients of instrumental meaning. *Journal of Semantics*, *34*, 507–537. https://doi.org/10.1093/jos/ffx003

Rissman, L., Rawlins, K., & Landau, B. (2015). Using instruments to understand argument structure: Evidence for gradient representation. *Cognition*, *142*, 266–290. https://doi.org/10.1016/j.cognition.2015.05.015

Rothstein, S. (2004). *Structuring events*. Malden, MA: Blackwell.

Sakarias, M., & Flecken, M. (2018). Keeping the result in sight and mind: General cognitive principles and language-specific influences in the perception and memory of resultative events. *Cognitive Science*, *43*, 1–30. https://doi.org/10.1111/cogs.12708

Scheepers, C., Hemforth, B., & Konieczny, L. (2000). Linking syntactic functions with thematic roles: Psych-verbs and the resolution of subject-object ambiguity. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing: Studies in theoretical psycholinguistics* (pp. 95–135). Dordrecht, the Netherlands: Springer.

Shipley, T. F., & Zacks, J. M. (2008). *Understanding events: From perception to action*. New York: Oxford University Press.

Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., & Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 5968–73. https://doi.org/10.1073/pnas.1423080112

Swallow, K., Zacks, J., & Abrams, R. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology*, *138*, 236–257. https://doi.org/10.1037/a0015631

Taylor, B. (1977). Tense and continuity. *Linguistics and Philosophy*, *1*, 199–220. https://doi.org/10.1007/BF00351103

ter Bekke, M., Özyürek, A., & Ünal, E. (2019). Speaking but not gesturing predicts motion event memory within and across languages. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 2940–2946). Montreal: Cognitive Science Society.

Ünal, E., & Papafragou, A. (2016). Interactions between language and mental representations. *Language Learning*, *66*, 554–580. https://doi.org/10.1111/lang.12188

Ünal, E., & Papafragou, A. (2019). How children identify events from visual experience. *Language, Learning and Development*, *15*, 138–156. https://doi.org/10.1080/15475441.2018.1544075

Ünal, E., Trueswell, J., & Papafragou, A. (2017). How children map event participants onto language. Paper presented at the symposium "Encoding events in language and cognition" (E. Ünal, Chair), 14th International Congress for the Study of Child Language, Lyon, France. Retrieved from osf.io/qhtms

van Hout, A., de Swart, H., & Verkuyl, H. J. (2005). Introducing perspective on aspect. In H. J. Verkuyl, H. de Swart, & A. van Hout (Eds.), *Perspectives on aspect* (pp. 1–17). Dordrecht, the Netherlands: Springer.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, *66*, 143–160. https://doi.org/10.2307/2182371

von Fintel, K., & Matthewson, L. (2008). Universals in semantics. *The Linguistic Review*, *25*, 139–201. https://doi.org/10.1515/TLIR.2008.004

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

Waxman, S., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development*, *1*, 139–156. https://doi.org/10.1016/S0885-2014(86)80016-8

Webb, A., Knott, A., & MacAskill, M. R. (2010). Eye movements during transitive action observation have sequential structure. *Acta Psychologica*, *133*, 51–56. https://doi.org/10.1016/j.actpsy.2009.09.001

Wilbur, R. (2003). Representation of telicity in ASL. *Chicago Linguistic Society*, *39*, 354–368.

Wilson, F., Papafragou, A., Bunger, A., & Trueswell, J. (2011). Rapid extraction of event participants in caused motion events. In L. Carlson, C. Hölscher, & T. F. Shipley (Eds.), *Proceedings from the 33rd Annual Meeting of the Cognitive Science Society* (pp. 1206–1211). Hillsdale, NJ: Erlbaum.

Wilson, F., Ünal, E., Trueswell, J., & Papafragou, A. (2014). Homologies between language and event cognition: Evidence from event role prominence. Paper presented at the 39th Annual Boston University Conference on Language Development, Boston, MA.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Wundt, W. (1970). The psychology of the sentence. In A. L. Blumenthal (trans.), *Language and psychology: Historical aspects of psycholinguistics* (pp. 20–31). New York: John Wiley & Sons.

Yin, J., & Csibra, G. (2015). Concept-based word learning in human infants. *Psychological Science*, *26*, 1316–1324. https://doi.org/10.1177/095679761558875

Zacks, J. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*, 979–1008. https://doi.org/10.1207/s15516709cog2806_5

Zacks, J., Speer, N., & Reynolds, J. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*, 307–327. https://doi.org/10.1037/a0015305

Zacks, J., Speer, N., Swallow, K., Braver, T., & Reynolds, J. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*, 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zacks, J., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3–21. https://doi.org/10.1037/0033-2909.127.1.3

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162. https://doi.org/10.1037/0033-2909.123.2.162

Zwitserlood, P., Bölte, J., Hofmann, R., Meier, C. C., & Dobel, C. (2018). Seeing for speaking: Semantic and lexical information provided by briefly presented, naturalistic action scenes. *PLoS ONE*, *13*(4), e0194762. https://doi.org/10.1371/journal.pone.0194762