UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Midpoints and Endpoints in Event Percept

Permalink
https://escholarship.org/uc/item/2jr4839f

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 40(0)

Authors
Ji, Yue
Papafragou, Anna

Publication Date
2018

Midpoints and Endpoints in Event Perception

Yue Ji (Jiyue@Udel.Edu)
Department of Linguistics & Cognitive Science, University of Delaware, 125 E. Main Street
Newark, DE 19716 USA

Anna Papafragou (Apapafragou@Udel.Edu)
Department of Psychological & Brain Sciences, University of Delaware, 105 The Green
Newark, DE 19716 USA

Abstract

Events unfold over time, i.e., they have a beginning and endpoint. Previous studies have illustrated the importance of endpoints for event perception and memory (Lakusta & Landau, 2005, 2012; Papafragou, 2010; Strickland & Keil, 2011; Zacks & Swallow, 2007). However, this work has not compared endpoints to other potentially salient points in the internal temporal profile of events (e.g., midpoints) and has only discussed events with a self-evident endpoint. In the present study, we explored sensitivity to event endpoints and midpoints in events of different types. Our results show that people are more disturbed by interruptions at the end compared to interruptions in the middle of an event – but only when perceiving a bounded event (i.e., an event with an inherent endpoint). This finding reveals complex tracking of the abstract internal temporal structure of events during event perception.

Keywords: endpoint, boundedness, event perception

Introduction

Events unfold over time and can be defined as temporal segments with "a beginning and an ending" (Zacks & Tversky, 2001). Much work has explored how viewers identify the beginning and ending of an event. People perceive an event boundary when significant changes in physical features such as direction, location or speed of action occur (e.g., a train arriving at the station; Zacks, Speer, Swallow, Braver, & Reynolds, 2007); more importantly, people encode event boundaries by making use of conceptual features such as goal-directedness, or causation (e.g., a researcher arriving at a conclusion; Zacks & Swallow, 2007).

Event boundaries such as the beginning and the ending of an event have a privileged status in memory and provide anchors for later learning and describing (Swallow, Zacks, & Abrams, 2009). Furthermore, event endpoints appear to be particularly important compared to event beginnings. For instance, many studies on motion events have revealed a source-goal asymmetry: the goal of motion (i.e., the endpoint, as in *Maria ran _to the store_*) is more accurately encoded in both language and memory as opposed to the source (as in *Maria ran _away from the house_*; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007; Wagner, 2009). Furthermore, the absence of an endpoint can trigger rapid reactions from event viewers. When the endpoint is not in sight, even infants as young as 12 months old can actively infer it from observed actions (Csibra, Biro, Koos, & Gergely, 2003). Relatedly, when the endpoint of a causally active

event within an event chain remains unseen, adults form inferences about this event within seconds. For instance, after watching videos of someone launching an object (e.g., kicking a soccer ball) followed by the object's directed motion (e.g., the ball flying into a goal), participants mistakenly reported that they saw the moment of contact at launching (i.e., the endpoint of the kicking event), even when it was actually omitted from the display (Strickland & Keil, 2011).

Previous studies on event structure leave open two issues. A first issue is that, apart from beginnings and endpoints, other potentially salient temporal landmarks such as midpoints have not received much attention. To our knowledge, only one recent study has compared event midpoints to both event starting points and endpoints (Gold, Zacks, & Flores, 2017). In this study, participants watched movies of everyday activities (e.g., setting up for a party) which were composed of a series of meaningful events (e.g., taking plates and napkins out of a bag, laying the table, etc.). Some of the movies were edited by placing cues (a bell sound along with an arrow pointing to the affected object) either at event midpoints or at event boundaries. It turned out that cues at event midpoints improved subsequent memory of movies, although the cues were less effective than cues at event boundaries. These findings suggest that event midpoints, although informative to some extent, are less salient compared to event starting points and endpoints.

A second issue is that past literature on event boundaries has focused on events that have well-defined endpoints. Nevertheless, for a broader range of events, endpoints may not be specified and notions of endpoint (and event boundaries more generally) may turn out to be very abstract. Inspired by the extensive linguistic literature on aspect (i.e., the linguistic encoding of the internal temporal profile of events; see Bach, 1986; Harley, 2003; Jackendoff, 1991), one could distinguish between two types of events with different internal structures and different ways in which they come to an end. *Bounded* events such as (1) have an inherent endpoint "which is there from the outset and culminates if not interrupted" (Mittwoch, 2013). In the example in (1), the endpoint is the moment when the last card is added to the pile. *Unbounded* events such as (2) lack an inherent endpoint and may terminate at any arbitrary moment. In the example in (2), the endpoint is not different from any other time slice of the event. Sentences encoding the two types of events can be modified by different temporal phrases. Bounded events go

with *in X time*, an interval adverbial specifying a delimited time span within which the endpoint has been achieved. Unbounded events can only be modified by *for X time*, a duration adverbial expressing how long the event has lasted.

(1)   The girl piled up the poker cards in / *for 1 minute.
(2)   The girl shuffled the poker cards *in / for 1 minute.

A recent study (Ji & Papafragou, 2017) showed that viewers can form categories corresponding to bounded and unbounded events after watching a series of video clips, even though they are better with bounded events (presumably because of the presence of defined endpoints that made such events 'atomic' and easier to compare and generalize over). The fact that viewers detect event boundedness reveals that event perception is sensitive to abstract considerations of event structure. This result strongly raises the possibility that the psychological representation of event boundedness has broader, testable implications about how people perceive the internal temporal structure of events (including midpoints as well as event boundaries) across different event types.

In the present study, we address this possibility. We build on observations about the internal composition of bounded and unbounded events that were initially formulated within linguistic theory. In this literature, bounded events are considered developments leading to a "built-in terminal point" (Comrie, 1976), "climax" (Vendler, 1957) or "culmination" (Parsons, 1990). If we divide a bounded event into temporal slices with minimal duration, each slice represents a different stage of development. For the bounded event in (1), if the girl's piling up of the poker cards takes one minute, then a one-second slice of the event is likely to be a distinct stage of the event (e.g., adding a card to the pile). By contrast, unbounded events are homogenous (Hinrichs, 1985) or cumulative (Krifka, 1989, 1998; Taylor, 1977). An unbounded event can be divided into any number of temporal slices and each slice can still be regarded as an event of the same kind. For the unbounded event in (2), if the girl's shuffling of the poker cards lasts for one minute, then each one-second slice of her action is still an event of shuffling the poker cards. This difference in internal homogeneity makes the bounded-unbounded distinction in the event domain reminiscent of the object-substance distinction in the object domain (with bounded events resembling 'atomic', structured objects such as a sandcastle and unbounded events resembling non-atomic, mass-like entities lacking internal structure such as sand; cf. Bach, 1986 for linguistic arguments).

We propose that these differences in abstract internal event structure should affect how viewers process and weigh temporal slices of different events. Since bounded events have a finely differentiated internal structure that is defined on the basis of the availability of an inherent endpoint, we expect that endpoints should be particularly salient over other event slices such as midpoints for such events. By contrast, since unbounded events have a largely undifferentiated internal structure, endpoints should be treated largely

similarly to midpoints or other points within the event's temporal profile. To test this prediction, we experimentally compare viewers' sensitivity to disruptions at different time points in the temporal profile of bounded vs. unbounded events. We expect that, for bounded events, disruptions during temporal endpoints should be dispreferred compared to disruptions during midpoints; for unbounded events, the endpoint-midpoint difference should disappear.

Experiment

Method

Participants One hundred and twenty adults (M_{age} = 19.6) participated in the experiment. All were undergraduates at the University of Delaware and received course credit for participation. Data from an additional group of 7 adults were collected but excluded due to experimenter error.

Stimuli Sixteen videos of bounded events and sixteen videos of closely related unbounded events were created (see Table 1). Related bounded-unbounded videos (i.e., each row in Table 1) had the same duration (range: 4.5s-13s; M = 8.6s) and involved the same actor wearing a yellow shirt. All of the videos began with the actor picking up an object or tool to perform an action, and came to an end with the actor putting down the object or tool and removing her hand. To create the videos, we were inspired by the linguistic literature detailing the factors that can determine whether an event is bounded or unbounded (see Filip, 2004; Tenny, 1987; cf. also Ji and Papafragou, 2017) and used two sources to create the contrast in boundedness across related events —the nature of the action and the nature of the affected object. For half of the cases, the bounded-unbounded events involved the same object but differed in terms of the nature of the action performed on the object: the bounded event displayed an action that caused a clear and temporally demarcated change of state in the object (e.g., fold up a handkerchief) while its unbounded counterpart did not involve such a change (e.g., wave a handkerchief). For the other half of the cases, the bounded and unbounded events involved the same action but differed in terms of the nature of the affected object: the bounded event involved a single object (e.g., draw a balloon) but its unbounded counterpart involved either an unspecified plurality of objects or a mass quantity (e.g., draw circles).

To ensure that all video stimuli would instantiate the contrast in boundedness, a separate group of 18 adults from the same population was asked to watch a subset of the clips and describe what happened in a full English sentence. The descriptions of stimulus events in Table 1 were the most frequent descriptions given by the participants. Participants' descriptions were tested for boundedness using the *in X time* versus *for X time* diagnostics. It turned out that the stimuli of bounded events elicited bounded descriptions 96.7 % of the time and the stimuli of unbounded events elicited unbounded descriptions 91.2 % of the time.

To increase the visual variety of the stimuli, we created an additional version of the videos that was identical to the first

Table 1: Stimulus events.

| Phase	Boundedness Source	No.	Bounded Events	No.	Unbounded Events
Training	Nature of Action	1	fold up a handkerchief	17	wave a handkerchief
		2	put up one's hair	18	scratch one's hair
		3	pile up a deck of cards	19	shuffle a deck of cards
		4	group pawns based on color	20	mix pawns of two colors
	Nature of Affected Object	5	draw a balloon	21	draw circles
		6	tie a knot	22	tie knots
		7	eat a pretzel	23	eat cheerios
		8	flip a postcard	24	flip pages
Testing	Nature of Action	9	dress a teddy bear	25	pat a teddy bear
		10	roll up a towel	26	twist a towel
		11	fill a glass with milk	27	shake a bottle of milk
		12	scoop up yogurt	28	stir yogurt
	Nature of Affected Object	13	peel a banana	29	crack peanuts
		14	blow a balloon	30	blow bubbles
		15	tear a paper towel	31	tear slices off paper towels
		16	paint a star	32	paint stuff
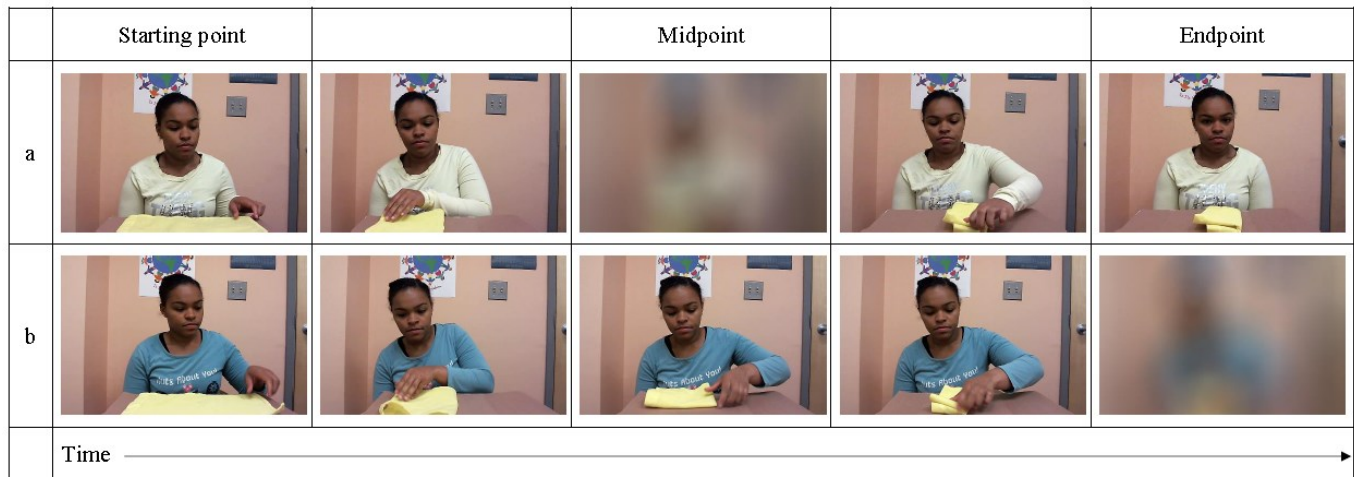
Figure 1. Examples of a training trial for a bounded event (folding up a handkerchief) that includes the two versions of the event: (a) mid-interruption (actor in yellow shirt), (b) end-interruption (actor in blue shirt).

except that the actor wore a blue shirt. All videos were then edited to introduce an interruption during which the screen turned blurry. Each video was edited twice, once to create a mid-interruption and once to create an end-interruption, depending on where the interruption was placed. In all cases, the interruption took up one-fifth of the total video duration.

Procedure We adopted a variant of the "picky puppet task" (Waxman & Gelman, 1986). Participants were invited to watch a couple of videos. The experimenter told participants that the girl in the videos liked performing, but was very picky about her videos: she liked some videos but not the others. The task was to figure out what kind of videos the picky girl liked.

Participants were randomly assigned to one of two conditions depending on the event type (bounded or unbounded) that they were exposed to throughout the experiment. In the training phase, participants watched a total of 16 videos. These were comprised of 8 events (No. 1-8 in Table 1 for bounded events, or 17-24 for unbounded events, presented in a random order), each with two versions shown in succession in the center of the screen. The two versions differed in terms of both the actor's shirt color (blue vs. yellow) and in terms of the placement of the interruption (mid-interruption vs. end-interruption - see Figure 1 for an

example). Within this phase, half of the time mid-interruptions occurred when the actor wore a blue shirt and the other half they occurred when the actor wore a yellow shirt. Even though our hypothesis targeted detection of a mid- vs. end-interruption, we added the change of shirt color to ensure that participants would treat the two (highly similar) versions of each event as different tokens. The order of mid-interruptions and end-interruptions, as well as shirt colors, was counterbalanced.

After each version the experimenter said either, "The girl likes the video", or "The girl doesn't like the video". Participants were randomly assigned to one of two conditions. In the "Likes mid-interruption" condition, they were always told that the picky girl liked the video after mid-interruptions but did not like the video after end-interruptions. In the "Likes end-interruption" condition, the girl's preference was reversed.

In the testing phase, participants watched a total of 8 videos corresponding to 8 new events (No. 9-16 in Table 1 for bounded events, or 25-32 for unbounded events, presented in random order). Half of these events were presented in their mid-interruption version and the other half in their end-interruption version. Participants were randomly assigned to one of two lists. Each list included one interruption version (mid- or end-) of each event; the agent's shirt color was kept constant for that event across lists. Types of interruptions and shirt color changes were evenly split within each list. After watching each video, participants were asked: "Will the girl like this video or not?" They were requested to give a *Yes/No* response on an answer sheet.

Results

A preliminary ANOVA performed on the proportion of correct responses with Source of Boundedness (i.e., Nature of Action vs. Nature of Affected Object) as a within-subjects factor revealed no significant effect of that factor (F (1, 119) = .569, p = .452). Therefore, answers to questions targeting the two sources of boundedness were collapsed for further analysis.

Results are shown in Figure 2. An ANOVA with the proportion of correct responses as the dependent variable, Condition (Likes mid-interruption vs. Likes end-interruption) and Event Type (Bounded vs. Unbounded) as between-subjects factors was performed. There was a significant effect of Condition (F (1, 116) = 6.26, p = .014): participants performed better when the picky girl liked mid-interruptions compared to end-interruptions. No effect of Event Type was detected (F (1, 116) = .70, p = .406): response accuracy did not differ across bounded and unbounded events. However, there was a significant interaction between Condition and Event Type (F (1, 116) = 7.12, p = .009). When participants watched videos of bounded events, they were better at identifying a preference for mid-interruptions compared to a preference for end-interruptions (t (38.24) = 3.39, p = .002). By contrast, when participants watched videos of unbounded events, they were equally good at identifying that the girl

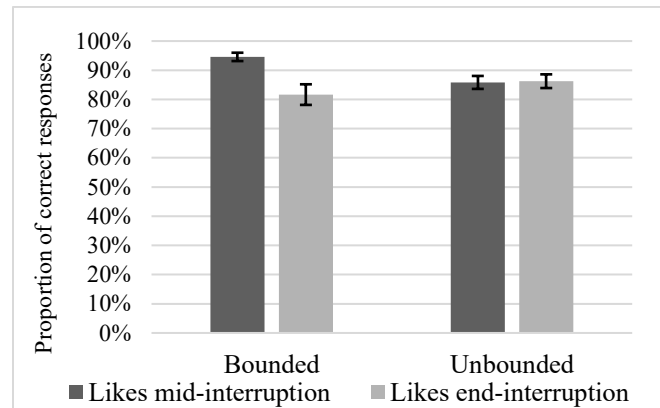liked mid-interruptions and end-interruptions (t (58) = -.13, p = .898).

Figure 2: Proportion of correct responses. Error bars represent standard error.

Discussion

For bounded (internally structured) events, blocking the endpoint was more disturbing to event viewers than blocking the midpoint (and hence it was harder to accept that the girl liked end- compared to mid-interruptions for such events). For unbounded (internally unstructured, homogeneous) events, there was no such difference. Together, our findings suggest that the salience of endpoints in event perception is tied to the internal structure of events.

General Discussion

The present experiment tested the hypothesis that people represent aspects of the internal temporal profile of events such as midpoints and endpoints differently depending on the perceived nature of the event (bounded vs. unbounded). For bounded events (i.e., those with an inherent endpoint and a finely differentiated structure), we expected that interruptions of the visual stimulus would be less disruptive is they appeared in the middle compared to the end of the event. For unbounded events (i.e., those lacking inherent endpoints and with a homogenous structure), we expected no difference between midpoint and endpoint interruptions. Our results confirmed these predictions. We conclude that people attend to abstract event structure in event perception. Furthermore, viewers process aspects of how an event develops differently depending on these abstract considerations.

Our findings provide new evidence for the importance of event endpoints, extending and enriching previous literature that compared endpoints with starting points in motion events (Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007; Wagner, 2009). More importantly, our results connect the salience of endpoints and other event slices to a broader framework concerning the fundamentals of event structure that has its roots in the linguistic literature. According to a widely shared assumption within this literature, "the notions 'bounded' and 'unbounded' belong to a finite set of primitives that characterizes parts of conceptual

structure." (Filip, 1993, p.10). Our results offer some of the first experimental demonstrations that event structure is indeed sensitive to such linguistically-inspired concepts (see also Ji & Papafragou, 2017). In this sense, even though our experiment did not ask participants to describe the events they viewed, and hence did not contain an overt linguistic task, its findings can be used to support the presence of parallels between event language and event perception (Folli & Harley, 2006; Lakusta & Landau, 2005; Malaia, 2014; Papafragou, 2015; Strickland et al., 2015; Tversky, Zacks, Morrison, & Hard, 2011).

The present data leave several directions open for further research. Our video stimuli have the clean setting of a lab room which highlights the actor, the action and the affected object with minimal information provided from the context. In reality, however, events occur in more informative and more complex contexts, which can influence whether an event is encoded as bounded or unbounded. For instance, in a setting where a couple of friends are about to play a poker game, shuffling the cards can be interpreted as a bounded event: it comes to an end when the cards are ready for the game. In this case, the inherent endpoint that defines a bounded event is provided by a salient intention (Depraetere, 2007). The boundedness of an event seems also to depend on one's knowledge about the larger scene (Filip, 2001). Future research needs to address how this type of higher-level knowledge works to constrain event construals.

Acknowledgments

References

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy, 9,* 5-16.

Barner, D., Wagner, L. & Snedeker, J. (2008). Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition, 106,* 805-832.

Comrie, B. (1976). *Aspect*: *An Introduction to the Study of Verb Aspect and Related Problems.* Cambridge: CUP.

Csibra, G., Biro, S., Koos, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27,* 111-133.

Depraetere, I. (2007). (A)telicity and intentionality. *Linguistics, 45,* 243-269.

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

Filip, H. (2001). Nominal and verbal semantic structure: analogies and interactions. *Language Sciences, 23,* 453-501.

Filip, H. (2004). The telicity parameter revisited. *SALT, 14,* 92-109.

Folli, R., & Harley, H. (2006). What language says about the psychology of events. *Trends in Cognitive Science, 10,* 91-92.

Gold, D., Zacks, J., & Flores, S. (2017). Effects of cues to event segmentation on subsequent memory. *Cognitive Research: Principles and Implications*, *2*, 1-15.

Harley, H. (2003). How do verbs get their names? Denominal verbs, manner incorporation and the ontology of verb roots in English. In N. Erteschik-Shir, & T. Rapoport (Eds.), *The Syntax of Aspect*. Oxford: OUP.

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

Jackendoff, R. (1991). Parts and boundaries. *Cognition, 41*, 9-45.

Ji, Y., & Papafragou, A. (2017). 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.

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, Groningen-Amsterdam Studies in Semantics 11*. Dordrecht: Foris Publications.

Krifka, M. (1998). The origins of telicity. In S. Rothstein (ed.), *Events and Grammar*. Dordrecht: Kluwer.

Lakusta, L., & Landau, B. (2005). Starting at the end: the importance of goals in spatial language. *Cognition, 96*, 1-33.

Lakusta, L., & Landau, B. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal. *Cognitive Science*, *36*, 517-544.

Malaia, E. (2014). It still isn't over: Event boundaries in language and perception. *Language and Linguistics Compass, 8*, 89-98.

Mittwoch, A. (2013). On the criteria for distinguishing accomplishments from activities, and two types of aspectual misfits. In B. Arsenijević, B. Gehrke & R. Marín (Eds.), *Studies in the Composition and Decomposition of Event Predicates*. Dordrecht: Springer.

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science, 34,* 1064-1092.

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*. Cambridge, MA: MIT Press.

Parsons, T. (1990). *Events in the Semantics of English*: *A Study in Subatomic Semantics*. Cambridge, MA: MIT Press.

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

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.

Strickland, B., & Keil, F. (2011). Event completion: Event

based inferences distorts memory in a matter of seconds. *Cognition, 121,* 409-415.

Swallow, K., Zacks, J., & Abrams, R. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology, 138*, 236-257.

Taylor, B. (1977). Tense and continuity. *Linguistics and Philosophy, 1,* 199-220.

Tenny, C. (1987). *Grammaticalizing Aspect and Affectedness* (Doctoral dissertation). MIT, Cambridge, MA.

Tversky, B., Zacks, J., Morrison, J., & Hard, B. (2011). Talking about events. In J. Bohnemeyer & E. Pederson (Eds.) *Event Representation in Language and Cognition*. Cambridge: Cambridge University Press.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review, 66*, 143-160. Reprinted as Chapter 4 of Z. Vendler (Ed.), *Linguistics in Philosophy* (1967).

Wagner, L. (2009). Manners and goals in pre-linguistic thought: the origins of aspectual construal. *Proceedings of the 33th annual BU Conference on Language Development* (pp. 599-610). Somerville: Cascadilla Press.

Waxman, S. & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*, 139-156.

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

Zacks, J., & Swallow, M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16,* 80-84.

Zacks, J., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3-21.