Situational Dimensions that drive Event Boundaries

Gabriel Kressin Palacios (gkressi1@jhu.edu)
Dhruva Arekar (darekar1@jh.edu)
Christopher J. Honey (chris.honey@jhu.edu)
Department of Psychological & Brain Sciences, Johns Hopkins University

Abstract

Humans maintain an internal representation of the agents, locations, goals, and causal relationships of their present situation. In the study of event cognition, two central ideas are that (i) people automatically segment continuous streams of experience into discrete event representations, and (ii) boundaries between events correspond to moments of prediction error from active event models. Here, we asked whether event boundaries could be reliably predicted directly from discontinuities in a small set of event dimensions, such as locations and goals within narrative text. We defined dimension rating criteria using event boundaries in an initial story and then applied these criteria to map discontinuities and predict event boundaries in a held-out natural text. Our decision tree model predicted the held-out event boundaries above chance. These preliminary results suggest that the detection of simple discontinuities, such as changes in location or agents within a narrative, provides a concrete and interpretable model of event boundary generation.

Keywords: event segmentation; narrative comprehension; event indexing model; event horizon model

Introduction

The ease with which we comprehend everyday experiences, such as a shopping trip, conceals a remarkable process by which we transform continuous sensory input into complex representations that describe ongoing states of affairs. According to Event Segmentation Theory, we construct such representations by segmenting continuous experiences into discrete events, with boundaries between events determined by transient increases in prediction error (Radvansky & Zacks, 2014, 2017). This theory has been successfully instantiated in computational models (Franklin, Norman, Ranganath, Zacks, & Gershman, 2020; Kumar et al., 2023). However, it is not immediately apparent which features of everyday experience drive the prediction error. Moreover, event segmentation theory has difficulty accounting for transitions between events within highly predictable sequences, such as one's daily coffee-making routine. Older theories of narrative comprehension emphasize the importance of causal relationships (Trabasso & Van Den Broek, 1985; Graesser, Singer, & Trabasso, 1994; Zwaan, Radvansky, Hilliard, & Curiel, 1998) and features such as agents and locations for constructing event representations. For instance, the Event-Indexing Model proposes that individuals monitor the continuity of events across five situational dimensions: space, time, causality, intentionality/goals, and protagonists/agents (Zwaan & Radvansky, 1998). When individuals encounter discontinuities in any dimension, they are thought to update their event representation. The importance of these dimensions is empirically supported by their effects on reading times (Zwaan, Magliano, & Graesser, 1995; Magliano, Miller, & Zwaan, 2001), brain activity (Reagh & Ranganath, 2023; Chen & Bornstein, 2024), and the correlation of these dimensions with increased probability of event segmentation in a movie (Zacks, 2010). However, the operationalization of discontinuities in situational dimensions is ambiguous and varies across studies. Therefore, we asked: to what extent do discontinuities in situational dimensions predict the location of event boundaries in real-world narratives?

Table 1: Definitions of Discontinuities

Dimension		Definition
New-Agent	$D_{\text{N-A}}$	Presence of a new character.
Protagonist	D_{P}	Change of the main acting agent(s). Patients do not count as protagonists. Agents are grouped if sensible, e.g during a conversation.
Location	D_{L}	Explicit or inferred change in location.
Time	D_{T}	Explicit or inferred change in time.
Goal	D_{G}	Inconsistency with a previous goal, or a new goal that is not a superordinate goal of a previous goal.
Causal	D_{C}	If sentence does not have a reasonable cause in previous text or is unexpected given the circumstances. A reasonable cause can be established using world and background knowledge.

Methods

Data

We used the stories "Pieman" (O'Grady, 2008) and "Tunnel Under the World" (Pohl, 1956) as narrative texts. "Pieman" ($\sim$ 7min, 936 words) was originally recorded for the Moth festival, while "Tunnel Under the World" (~ 25min, 3,432 words) was broadcast on the radio show "X Minus One". We used the transcripts and consensus boundaries provided by Michelmann, Kumar, Norman, and Toneva (2025), which were obtained by smoothing and thresholding boundary ratings recorded during auditory presentation of the stories. Boundary data for "Pieman" were collected from 205 online participants (Michelmann et al., 2021); boundary data for "Tunnel Under the World" from 10 in-person participants (Lositsky et al., 2016).

Table 2: Sentences, discontinuities, and consensus boundaries ("Tunnel Under the World"). Presence of a discontinuity and a boundary are marked by a 1.

Sentence	$D_{\text{New-A}}$	D_{P}	D_{L}	D_{T}	D_{G}	D_{C}	B
GUY: I'll see you tonight.	0	0	0	0	1	0	1
NARRATOR: Guy Burckhardt got on his bus.	0	0	1	1	0	0	1
NARRATOR: There were the same unfamiliar faces, the same unusually new looking buildings, the same unusually bright sunshine.	0	0	0	0	0	0	1
NARRATOR: And on the customary corner, Henry Swanson, pale and furtive, climbed aboard.	0	1	0	0	0	0	0

Rating Discontinuities.

Although definitions for dimensional discontinuities exist (Zwaan et al., 1998; Zwaan, Magliano, & Graesser, 1995; Magliano et al., 2001; Zwaan, Langston, & Graesser, 1995), their application to real-world narratives contains ambiguities; therefore, we developed our own set of instructions. We used "Pieman" to develop discontinuity definitions (see Table 1) that aligned with human consensus boundaries. Note that these definitions are yet far from perfect and still leave considerable room for interpretation, an issue we hope to address in the future work. After developing the rating instructions, two raters independently rated "Tunnel Under the World". Table 2 shows the discontinuity ratings and the event boundaries for three example sentences.

Computing Event Boundaries

To predict event boundaries, we used 5-fold cross-validation to fit a decision tree from the dimension ratings to human consensus boundaries. During fitting, samples were reweighted to account for the large number of sentences without consensus human boundaries. For each fold, we computed the Hamming distance between predicted and observed boundaries (lower is better; see Michelmann et al. (2025)), as well as precision, recall, and the correlation between binary vectors of predicted and observed boundaries. We generated null distributions for each test statistic by shuffling the order of event boundaries between events within each cross-validation fold, with significance determined by the 2.5th-97.5th percentile range [1].

Results

The model predicted the location of event boundaries above chance in the held-out story "Tunnel Under the World". For the first rater, averaging across all folds, we obtained a Hamming distance of $.11$ (null 2.5th-97.5th percentiles: $.127$, $.156$), precision of $.21$ (null percentiles: $.0$, $.13$), recall of $.51$ (null percentiles: $.0$, $.373$), and a correlation of $.26$ (null percentiles: $-.079$, $.141$). On average, the decision tree predicted 9.8 boundaries compared to the 5.4 human boundaries present

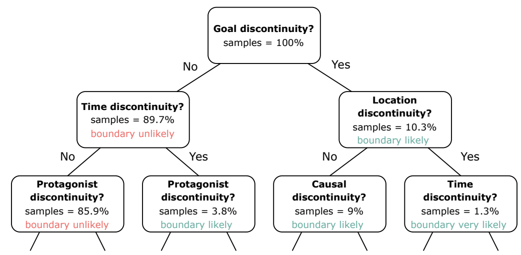

Figure 1: Truncated decision tree of the fold with the best Hamming distance. The goal dimension being at the root note indicates that it has strong predictive power.

in each fold. Figure 1 shows the decision tree from the fold with the shortest Hamming distance. The model predictions based on the second rater were significant for all metrics except precision[2]. We obtained a Hamming distance of $.08$ (null 2.5th-97.5th percentiles: $.082$, $.103$), precision of $.3$ (null percentiles: $.0$, $.333$), recall of $.36$ (null percentiles: $.0$, $.333$), and a correlation of $.26$ (null percentiles: $-.049$, $.246$). The average number of model predictions was 4.6 compared to 5.4 human boundaries.

Discussion

We found that discontinuities in dimensions such as agents and location provide a simple and interpretable model of event segmentation. Thus, it may be fruitful to update current models of event segmentation to incorporate the dimension-based approach from theories such as the Event-Indexing Model. In the future, we plan to refine the discontinuity definitions, determine which combinations of dimensions are most important for accurate predictions, and measure the generalizability of our approach across stories with different styles (e.g., podcast-like style such as "Monkey in the Middle"). Finally, we plan to directly compare our dimensional model against models from the prediction error framework (Kumar et al., 2023).

[1] The code and data are available online: https://github.com/GabrielKP/ccn25-dimensions-boundaries

[2] We did not have causal dimension ratings for the second rater.

Acknowledgments

We gratefully acknowledge the support of NSF CAREER Award 2238711.

References

Chen, J., & Bornstein, A. M. (2024, August). The causal structure and computational value of narratives. *Trends in Cognitive Sciences*, *28*(8), 769–781. Retrieved 2025-02-08, from https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(24)00082-2 (Publisher: Elsevier) doi: 10.1016/j.tics.2024.04.003

Franklin, N. T., Norman, K. A., Ranganath, C., Zacks, J. M., & Gershman, S. J. (2020, March). Structured Event Memory: A neuro-symbolic model of event cognition. *Psychological Review*, *127*(3), 327. Retrieved 2023-06-13, from https://psycnet.apa.org/fulltext/2020-21913-002.pdf (Publisher: US: American Psychological Association) doi: 10.1037/rev0000177

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. Retrieved 2024-04-03, from https://doi.apa.org/doi/10.1037/0033-295X.101.3.371 doi: 10.1037/0033-295X.101.3.371

Kumar, M., Goldstein, A., Michelmann, S., Zacks, J. M., Hasson, U., & Norman, K. A. (2023). Bayesian Surprise Predicts Human Event Segmentation in Story Listening. *Cognitive Science*, *47*(10), e13343. Retrieved 2024-06-13, from https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.13343 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.13343) doi: 10.1111/cogs.13343

Lositsky, O., Chen, J., Toker, D., Honey, C. J., Shvartsman, M., Poppenk, J. L., . . . Norman, K. A. (2016, November). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *eLife*, *5*, e16070. Retrieved 2025-04-11, from https://doi.org/10.7554/eLife.16070 (Publisher: eLife Sciences Publications, Ltd) doi: 10.7554/eLife.16070

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001, September). Indexing space and time in film understanding. *Applied Cognitive Psychology*, *15*(5), 533–545. Retrieved 2024-04-03, from https://onlinelibrary.wiley.com/doi/10.1002/acp.724 doi: 10.1002/acp.724

Michelmann, S., Kumar, M., Norman, K. A., & Toneva, M. (2025, January). Large language models can segment narrative events similarly to humans. *Behavior Research Methods*, *57*(1), 39. Retrieved 2025-03-13, from https://doi.org/10.3758/s13428-024-02569-z doi: 10.3758/s13428-024-02569-z

Michelmann, S., Price, A. R., Aubrey, B., Strauss, C. K., Doyle, W. K., Friedman, D., . . . Norman, K. A. (2021, September). Moment-by-moment tracking of naturalistic learning and its underlying hippocampo-cortical interactions. *Nature Communications*, *12*(1), 5394. Retrieved 2025-04-11, from https://www.nature.com/articles/s41467-021-25376-y (Publisher: Nature Publishing Group) doi: 10.1038/s41467-021-25376-y

O'Grady, J. (2008). *Pie Man: The Moth (en-US) | Stories | Pie Man.* Retrieved 2025-04-11, from https://themoth.org/stories/pie-man

Pohl, F. (1956). *Tunnel Under the World | X Minus One.* Retrieved 2025-04-11, from https://www.genericradio.com/show/2CCNTKCBO

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

Radvansky, G. A., & Zacks, J. M. (2017, October). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, *17*, 133–140. Retrieved 2024-10-21, from https://www.sciencedirect.com/science/article/pii/S2352154617300037 doi: 10.1016/j.cobeha.2017.08.006

Reagh, Z. M., & Ranganath, C. (2023, March). Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. *Nature Communications*, *14*(1), 1279. Retrieved 2023-06-11, from https://www.nature.com/articles/s41467-023-36805-5 (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/s41467-023-36805-5

Trabasso, T., & Van Den Broek, P. (1985, October). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*(5), 612–630. Retrieved 2024-04-03, from https://linkinghub.elsevier.com/retrieve/pii/0749596X8590049X doi: 10.1016/0749-596X(85)90049-X

Zacks, J. M. (2010). The brain's cutting-room floor: segmentation of narrative cinema. *Frontiers in Human Neuroscience*, *4*. Retrieved 2025-01-16, from http://journal.frontiersin.org/article/10.3389/fnhum.2010.00168/abstract doi: 10.3389/fnhum.2010.00168

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995, September). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, *6*(5), 292–297. Retrieved 2023-02-20, from http://journals.sagepub.com/doi/10.1111/j.1467-9280.1995.tb00513.x doi: 10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 386–397. (Place: US Publisher: American Psychological Association) doi: 10.1037/0278-7393.21.2.386

Zwaan, R. A., & Radvansky, G. A. (1998). Situation Models in Language Comprehension and Memory. *Psychological bulletin*, *123*(2), 162–185.

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998, July). Constructing Multidimensional Situation Models During Reading. *Scientific Studies of Reading*, *2*(3), 199–220. Retrieved

2024-04-03, from http://www.tandfonline.com/doi/abs/10.1207/s1532799xssr0203_2 doi: 10.1207/s1532799xssr0203_2