Title
Reading the Language of Action

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN
1069-7977

Authors
Hard, Bridgette Martin
Recchia, Gabriel

Publication Date
2006

Peer reviewed

Reading the Language of Action

Bridgette Martin Hard (hard@psych.stanford.edu)
Gabriel Recchia (grecchia@stanford.edu)
Department of Psychology, Stanford University
Bldg. 420, Jordan Hall, Stanford, CA 94305 USA

Abstract

According to the *hierarchical encoding hypothesis*, people segment observed behavior into units of action that are integrated into larger units across time. Study 1 tested this hypothesis by adapting techniques from studies of reading: participants watched a self-paced slideshow of an activity. Participants looked longer at slides corresponding to boundaries between action units than at slides within units. This time may reflect integration of information within that unit before processing the next. Supporting this possibility, time spent at boundaries between units increased more for coarse units of action than for fine units of action. Studies 2 and 3 suggested other reasons that looking time increases at unit boundaries: boundaries between actions correspond to noticeable discontinuities in the flow of action and are inherently more interesting to observers than other parts of an action sequence.

Keywords: action parsing; hierarchical encoding; event segmentation

Introduction

Everyday interactions—from buying a cup of coffee to playing a game of basketball—depend on the ability to make sense of others' behavior. This ability is remarkable given that human behavior is so dynamic, presenting observers with a continuous, multi-media stream of physical information. Part of understanding this stream is to *segment* it into distinct actions like "pouring milk" or "shooting a basket" (Baldwin & Baird, 1999; Newtson, 1973; Zacks, Tversky, & Iyer, 2001). How is this accomplished? Part of the answer is that people infer and monitor the intentions of the actor (Baird, Baldwin, & Malle, 2006; Zacks et al, 2001). For example, when asked to view films of behavior and to press a key to mark off boundaries or *breakpoints* between individual actions (Newtson, 1973), people identify breakpoints that correspond to changes in intentions, cued by changes in the states of objects in a scene and also by predictable changes in the actor's body (Baldwin & Baird, 1999; Newtson, Engquist, & Bois, 1977).

The ability to make sense of behavior is also remarkable given the complexity of everyday activities: they can be segmented in different ways and on different levels. This is because the underlying structure of most activities is a *hierarchy* of parts and sub-parts, reflecting the hierarchical structure of the actor's goals and subgoals (e.g., Abbott, Black, & Smith, 1985; Bower, Black, & Turner, 1979; Newell & Simon, 1972; Schank & Abelson, 1977). Thus, activities can be segmented on a *coarse* temporal grain, corresponding to larger goals, or on a *fine* temporal grain, corresponding to component subgoals.

Given the hierarchical organization in behavior, at what level do people segment it and why? One solution proposed to this problem is that people select a level on which to segment behavior based on how much information they need to gain. Segmenting behavior more finely or more coarsely allows observers to gain more or less information, respectively (e.g., Newtson, 1973).

A different possibility is that people segment on a single level in response to instructions, but they typically attend to and segment behavior on many levels simultaneously (Hard, Lozano, & Tversky, in press; Zacks et al., 2001). This possibility makes intuitive sense given how people process other types of multi-level information streams. Understanding discourse, for example, requires identifying units of meaning on many levels: morphemes, words, phrase constituents, and so on. Perhaps people read observed behavior similarly to the way they read a book: by identifying and integrating meaningful units across time, likely based on their goal-subgoal relations. This view of action understanding can be called the *hierarchical encoding hypothesis*. The present studies aimed to test this view.

Study 1: Evidence of Hierarchical Encoding in On-Line Processing?

Theories of on-line discourse processing propose that people construct a representation of linguistic information that is updated whenever the listener or reader reaches the end of a unit (e.g., Gernsbacher, 1985; van Dijk & Kintsch, 1983). Support for these theories comes from studies using a *moving-window method*, in which readers see one word at a time and press a key to reveal the next word. This technique allows experimenters to measure reading time for individual words as an index of processing load. Findings across several studies have shown that reading time is longer for words located at major clause boundaries and even longer at sentence boundaries compared to non-boundaries (Haberlandt & Graesser, 1989; Haberlandt, Graesser, Schneider & Kiely, 1986). Increased reading time at clause boundaries and sentence boundaries is often referred to as "wrap-up time" and is believed to reflect demands of interpreting and integrating information within and across units of text. Wrap-up time is greater for sentence boundaries because more information must be integrated across a sentence than across a clause. Wrap-up time appears to be functional: text recall performance is greater when reading times at sentence-final words are longer.

If people understand observed behavior hierarchically, in the same way they understand discourse, then similar patterns might be observed in on-line processing of an action sequence. The present study tested this idea by adapting the moving-window method. Observers viewed a self-paced slideshow of an everyday behavior. The looking time for each slide served as an index of processing load. If action is understood by segmenting and integrating units across time, then looking times may show a "wrap-up" effect, such that slides at breakpoints between units are looked at longer than slides within units. Compared to slides at fine unit breakpoints, slides at coarse unit breakpoints should be looked at longer because they require many smaller units to be integrated and redescribed as a larger one.

Method

Forty participants first viewed a slideshow of one of four activities (*cleaning a dorm room, eating breakfast, building a TV cart*, or *putting on makeup*). The slideshows were constructed by taking a 3min long filmed activity, and extracting 1 frame from each 1s interval. Slideshow viewing was self-paced, and looking times for each slide were measured. Participants studied the actions for later recall, and then completed a free recall test. Because behavior, unlike text, lacks punctuation marks, different viewers might segment and organize a behavior stream in different ways. To account for this possibility, each participant was asked to identify levels of organization within the behavior stream after they completed the recall test. Participants identified levels of organization by viewing the film version of the activity and pressing a key to segment what they perceived as separate actions. Participants segmented the film 3 times, into *fine*, *intermediate*, and *coarse*-sized units. Half segmented small units first, and increased unit size on each subsequent viewing. Half did the opposite.

Results and Discussion

Does Looking Time Increase at Breakpoints? Based on the idea that observers update their representation of an action sequence at breakpoints, observers should spend more time processing slides at unit breakpoints than slides within units. Looking time data for each participant was log transformed to reduce positive skewness. There was a trend for participants to look longer at slides presented earlier in the slideshow. This trend was well-described by a power function. Each participant's looking time data was thus fitted with a power function using a curve estimation regression technique to factor out any influence of this power function in further analyses. Fitting power functions to the looking time data for each participant accounted for an average 48% of variance (SEM = 3%). All further analyses were conducted on the residuals from this regression analysis. Because residuals represent the difference between actual looking time for a given slide and the looking time predicted by the corresponding power function, the values of these residuals can be negative.

Recall that each slide in the slideshow represented 1s of the video version of the same activity. For each participant, the looking time for each slide was thus binned as a *unit breakpoint* if the participant indicated a breakpoint in the same 1s interval as that slide when segmenting the video. A slide was binned as *within-unit* if the participant indicated no breakpoint within that interval. The means and standard deviations were then calculated for each bin. A looking time score for each bin was calculated by dividing the mean looking time for that bin by its standard deviation. This looking time score represents an effect size estimate: how much the mean looking time in a given bin differs from the average looking time for all slides.

As predicted, participants looked longer at *unit breakpoints* (M = 0.18, SEM = 0.03) than at points *within-units* (M = -0.02, SEM = 0.01), *paired-t*(39) = 5.62, $p < .001$. This supports the hypothesis that observers wrap up at breakpoints.

Does Looking Time Increase with Unit Size? Observers should need more time to integrate larger quantities of information, leading to longer looking times for *coarse* breakpoints than for *intermediate* or *fine* breakpoints. To test this hypothesis, looking times already binned as *unit breakpoints* were further binned as *fine*, *intermediate*, or *coarse* breakpoints. On average, participants identified 53.15 (SEM = 8.01) *fine* breakpoints, 22.10 (SEM = 3.88) *intermediate* breakpoints, and 8.90 (SEM = 1.46) *coarse* breakpoints.

The results are shown in Figure 1. Unit size significantly affected looking time scores, $F(1.7, 64.8) = 5.43$, $p < .01$. As predicted, looking time increased linearly as a function of unit size: looking times scores were highest for *coarse* breakpoints (M = 0.35, SEM = 0.09), followed by *intermediate* breakpoints (M = 0.22, SEM = 0.05), and then *fine* breakpoints (M = 0.11, SEM = 0.03), $F(1, 39) = 8.23$, $p < .01$. Because there were no effects or interactions of activity or segmentation order, the data were collapsed across conditions. A Greenhouse-Geisser correction was applied to the degrees of freedom to adjust for violation of sphericity.

Does Looking Time Predict Recall? Do people really look longest at *coarse* breakpoints because they are integrating smaller units into larger ones, as the hierarchical encoding hypothesis proposes? If so, then we can make a prediction: integration and chunking of information into larger units is an effective memory strategy (e.g., Miller, 1956; Tulving, 1962); if looking time at coarse breakpoints reflects integration of fine-level actions into coarser ones, then it should predict better recall of those fine-level actions. Participants' free recall of the slideshow was coded by simply counting the number of actions reported. Participants recalled an average 23.43 actions (SEM = 1.62), most of them very fine-level, like "she took 3 bites of a banana." One participant was excluded from the present analysis as an outlier for recalling 51 actions—more than 2.5 standard deviations above the mean. Participants with longer looking time scores at *coarse* breakpoints recalled significantly more actions overall, $r(38) = .46$, $p < .01$. Looking time scores at other breakpoint levels did not predict recall. This result is consistent with the proposal that processing time at coarse breakpoints reflects integration of smaller units into larger chunks.

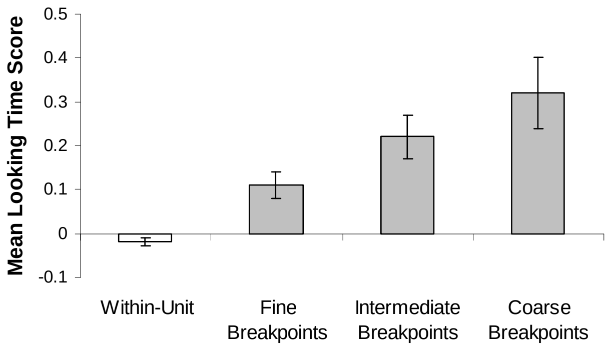

Figure 1: Mean looking time scores for slides that fell *within* action units, or at *fine*, *intermediate*, or *coarse*-unit breakpoints.

Does Looking Time Predict Enclosure? When observers view a film of behavior and divide it into fine and coarse units on two separate viewings, there is a reliable tendency for coarse breakpoints to fall close to, but slightly *after*, a fine breakpoint in time. This phenomenon is called *enclosure* (Hard et al., in press). The enclosure phenomenon appears to reflect an underlying, hierarchical encoding of observed behavior. *Enclosure scores,* based on the proportion of coarse-level breakpoints that fit this pattern, correlate with observers' spontaneous descriptions of hierarchical organization in behavior, across several studies (Lozano, Hard, & Tversky, in press; Hard et al, in press). Enclosure scores also increase reliably when observers are explicitly looking for coarse and fine units that are hierarchically related (Hard et al., in press).

One explanation for enclosure is that it results from the demands of integration processes at coarse breakpoints. Increased processing load might slow down observers in marking off a coarse breakpoint relative to a fine breakpoint, producing a segmentation pattern where coarse breakpoints fall close to, but slightly after fine breakpoints. Supporting this interpretation, participants in the present study with longer looking time scores at *coarse* breakpoints had higher mean enclosure scores, $r(39) = .45$, $p < .01$. Mean enclosure scores were ($M = .61$, $SEM = .02$), significantly different from a chance enclosure score of .5, $t(39) = 5.87$, $p < .001$, consistent with previous findings.

In Sum As observers watched a slideshow of an everyday activity, processing load increased at breakpoints between actions. Processing time increased more if those breakpoints corresponded to moments when several smaller action units could be integrated and chunked into a single unit. The more units that could be integrated, the more processing load increased. These increases in processing load appeared to reflect integration of smaller units into larger chunks, because participants with the greatest increases were able to recall more actions from the sequence. Participants with the greatest increases in processing time were also more likely to show *enclosure* in later patterns of segmentation. In previous studies, enclosure patterns correlated with people's descriptions of the hierarchical relations among a set of actions (Lozano et al., in press; Hard et al., in press). Together, these findings support the hierarchical encoding hypothesis: that people understand observed behaviors by segmenting and organizing them into hierarchically-organized action units.

Study 2: Hierarchical Encoding or Change Detection?

An important feature of breakpoints is that they correspond to discontinuities in the behavior stream—often moments when an actor dramatically reconfigures her body as she completes one goal and begins pursuing the next (Newtson et al., 1977). These reconfigurations likely involve a lot of movement, which means that slides at and near breakpoints might look very dissimilar from one another, at least relative to slides within units. Thus, observers might look longer at breakpoint slides because they are detecting dissimilarity, or a change between that slide and the immediately previous slide.

It is also possible that reconfigurations of the actor's body are more dramatic as unit size increases from fine to coarse. If so, then people might look longer at coarse breakpoints than intermediate and fine breakpoints because of greater amounts of change at these breakpoints, and not because they are integrating smaller action units into larger wholes. Study 2 tests this alternative account of the results from Study 1, through an analysis of change across an action sequence and its relationship to segmentation and looking time.

Method

The present study examined the amount of change occurring in the slideshows described in Study 1. Each frame of the slideshow was first passed through a convolution filter that identifies high contrast areas of the image (e.g., near the edges of people and objects in each frame). This process eliminates irrelevant data and restricts later calculations of change to the most salient features of the image. Each resulting image was then inverted and each pixel was assigned a numeric "contrast value" from 0 to 255 corresponding to its brightness, as defined by the Hue-Saturation-Brightness color model. Contrast values were high for pixels near the edges of people and objects in each frame, but low elsewhere.

A change index for each slide was calculated by pairing each pixel from that slide with the corresponding pixel in the previous slide, calculating the absolute value of the difference between the contrast values of every such pair, and summing up these absolute values. If the actor's body is moving and reorienting dramatically from one slide to another, then the contrast values of corresponding pixels in the two slides should be more different than if there is little movement and reorientation. For ease of interpretation, change indexes were standardized for each slideshow. The *makeup* slideshow was excluded from subsequent analyses because several random lighting changes in the slides made the change calculations unreliable.

Results and Discussion

Does Change Increase at Breakpoints? Do breakpoints correspond to relatively high amounts of change? For each participant from Study 1, the change index for each slide was binned as a *unit breakpoint* or as *within-unit*, using the same technique that was used to bin looking times. The means and standard deviations were then calculated for each bin. The mean change index for each bin was then adjusted by dividing it by its standard deviation.

Breakpoint slides corresponded to greater amounts of change (M = 0.24, SEM = 0.04) than slides *within-units* (M = -0.14, SEM = 0.03), *paired-t*(29) = 7.25, $p < .001$. This is consistent with previous findings that breakpoints involve relatively dramatic bodily reorientations by the actor (Newtson et al., 1977).

Does Change Increase with Unit Size? As shown in Figure 2, change was affected by unit size, $F(1.24, 36.06)$ = 9.28, $p < .01$. There was a significant linear trend, such that *coarse breakpoints* had the highest change index (M = 0.71, SEM = 0.16), followed by *intermediate breakpoints* (M = 0.25, SEM = 0.05), and then *fine breakpoints* (M = 0.19, SEM = 0.04), $F(1, 29)$ = 11.38, $p < .01$.

Does Change at Breakpoints Account for Looking Time Effects? Physical changes in the action sequences, like looking time, increased at unit breakpoints, and increased more as unit size got coarser. Given the similar pattern of results for change and looking time, it seems likely that change and looking time are related. Indeed, the change index for each slide correlated with looking times for 20 of the 30 participants. These correlations were weak (M = .19, SEM = .04), but were significantly different from 0, *one sample* $t(29)$ = 4.77, $p < .001$.

Given these findings, it is possible that participants in Study 1 looked preferentially at breakpoints because they were detecting this change, and not because they were segmenting and organizing the behavior stream at breakpoints. Suggesting against this possibility, breakpoints predicted looking time, even controlling for change. This was determined with a regression analysis in which each slide in the three slideshows *breakfast*, *cleaning*, and *TV cart* was treated as a "subject." For each slide, three values were calculated. First, the looking times of the 10 participants who viewed a slide were pooled and averaged to determine the mean looking time in ms. Mean looking times for each slideshow were then transformed with a $\log_{10}$ function and de-trended by fitting a power function and extracting residuals. Second, the number of times that a given slide was selected as a *coarse*, *intermediate*, or *fine* breakpoint was calculated. The logic here was that if breakpoints affect looking time for individual participants, then for the *group* of participants, the number of times a slide was selected as a breakpoint should affect the mean looking time. Third, the change index for that slide relative to the previous slide was calculated. All calculations just described were then standardized for each film.

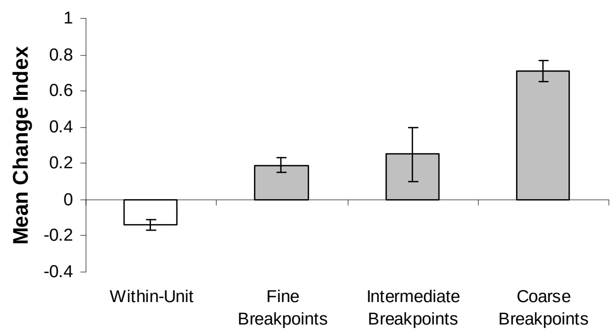

Figure 2: Mean change index for slides that fell *within* action units, or at *fine*, *intermediate*, or *coarse*-unit breakpoints.

The numbers of *intermediate* and *coarse* breakpoints in a given 1s interval were significant predictors of mean looking time, even controlling for change (for *intermediate*, B = 0.09, SE = 0.04, $t(1)$ = 2.62, $p < .001$; for *coarse*, B = 0.20, SE = 0.04, $t(1)$ = 5.45, $p < .001$). The number of *fine* breakpoints in a given 1s interval did not significantly predict looking time at the group level.

Second, according to regression coefficients, *coarse* breakpoints led to larger increases in looking time than *intermediate* breakpoints, even controlling for change. Specifically, if the number of *coarse* breakpoints selected for a given slide increased by 1 standard deviation, then the mean looking time for that slide increased by 0.20 of a standard deviation. Increasing the number of *intermediate* breakpoints selected for a given slide only increased mean looking time by 0.09 of a standard deviation. Although these regression coefficients offer only a rough approximation of how much *intermediate* and *coarse* breakpoints influence looking time, their values suggest that participants looked longer at *coarse* breakpoints than at *intermediate* breakpoints, even controlling for physical changes between slides.

In Sum The present study tested an alternative interpretation of results from Study 1: that breakpoints correspond to more change as the actor orients from one goal to the next, and that observers look longer at breakpoints because they are detecting this change. Results indicated that breakpoints do, in fact, correspond to relatively high amounts of change, and that this change becomes greater as units increase in size from fine to coarse. Both of these findings are consistent with previous research (Hard et al., in press; Newtson et al., 1977), but this is the first study to demonstrate that different levels of organization in human behavior correspond to changes of different magnitudes. This finding has interesting implications: it suggests physical “cues” in the behavior stream that might assist observers in identifying hierarchical structure.

Despite the relationship between change and breakpoints, change could not fully account for results from Study 1. Thus, it remains likely that patterns of looking time across an action sequence reflect changes in processing load as observers segment and organize observed actions into a hierarchical representation.

Study 3: Hierarchical Encoding or Informativeness?

Previous research indicates that breakpoints in action convey a lot of the meaning of an action sequence. Looking at a comic strip composed of breakpoints from an action sequence can be almost as comprehensible as watching the action sequence itself (Newtson & Engquist, 1976). Why would sequences of breakpoints be so informative? Recall that breakpoints typically correspond to moments when an actor completes one goal and begins to pursue the next (Baird et al., 2006). In activities that involve objects, completing a goal typically means a distinctive, meaningful change in the object or objects being acted on—changes that can define an action and confirm observers’ understanding of the action they’ve just seen. As the actor completes a goal, shifts in eye gaze and body direction might cue observers as to the next relevant goal, such as a new set of objects to be acted on. These features of breakpoints might explain why breakpoints can convey the gist of an action sequence, and why they tend to be better recognized than other points within the behavior stream (Newtson & Engquist, 1976).

Thus, observers in Study 1 might have looked longer at breakpoints because they represent information-rich parts of the behavior stream. The present study tests this possibility by showing a group of participants the slideshows from Study 1 in a scrambled order. If breakpoints are inherently more informative than other parts of an action sequence, then observers should look longer at slides later identified as breakpoints. Furthermore, looking at these breakpoints slides should predict how much information observers get from the action sequence.

Method

The methods used in the present study were identical to the methods used in Study 1, except that slides for a given activity were presented in a randomized order. Forty participants viewed one of the four activities, and then recalled actions from it. Participants then performed the segmentation task on the video version of the slideshow, shown in its original form.

Results and Discussion

Looking time data were again log transformed and de-trended with power functions. Fitting power functions to the looking time data for each participant accounted for an average 43% of variance (*SEM* = 3%). Participants looked longer at *unit breakpoints* ($M = 0.07$, $SEM = 0.03$) than at slides *within-units* ($M = -0.04$, $SEM = 0.04$). Although this difference was not as striking as for Study 1, it was significant, *paired-t*(39) = 3.40, $p < .01$. As can be seen in Figure 3, looking time for breakpoints did not increase as unit size increased from fine to coarse, $F(1.03, 40.24) < 1$, *ns*. Thus, effects of unit size on looking time in Study 1 don’t seem explainable by differences in how informative coarse breakpoints are relative to intermediate and fine breakpoints.

If breakpoints convey meaningful information about the behavior stream, observers who selectively look at breakpoints should be able to make the best inferences about what is happening in the slideshow and thus should be able to list more actions later. Disconfirming this prediction, recall was uncorrelated with looking time scores at *coarse breakpoints*, $r(38) = .15$, $p = .35$, *intermediate breakpoints*, $r(38) = .10$, $p = .53$, *fine breakpoints*, $r(38) = .13$, $p = .44$, or *breakpoints* overall, $r(38) = .19$, $p = .24$. Participants recalled an average 14.58 actions (*SEM* = 0.95), significantly less than participants in Study 1, $t(78) = -4.71$, $p < .001$.

In Sum The present study tested whether processing load increases at breakpoints because they are informative parts of an action sequence. People did in fact look longer at slides corresponding to breakpoints, even when slides were presented out of sequence. This suggests that breakpoints are inherently more interesting than other parts of an action sequence. Looking times for breakpoints in this study where much shorter than in Study 1, however, and looking times did not increase as unit size increased. Furthermore, looking time at breakpoints did not predict amount of later recall, suggesting that although breakpoints were interesting to look at, spending more time looking at them did not help participants to understand and remember what they saw. It remains plausible, however, that breakpoints are truly informative, but observers need to see them in a natural, sequentially presented action sequence to make use of their information content.

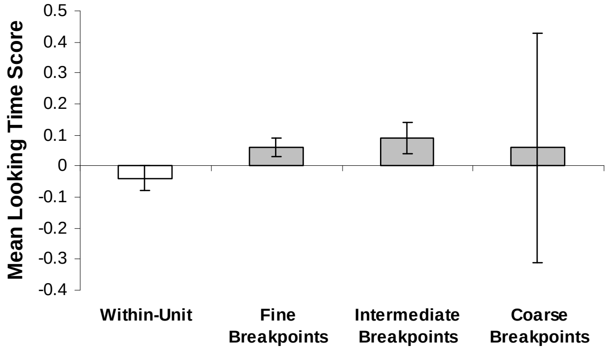

Figure 3: Mean looking time scores for slides in Study 3 that fell within action units, or at fine, intermediate, or coarse-unit breakpoints.

General Discussion

How do people make sense of the continuous and complex behaviors that other people perform? The hypothesis explored here is that people encode observed behavior as a set of hierarchically organized units. This involves identifying and integrating units across time, usually based on goal-subgoal relations. Supporting this *hierarchical encoding* hypothesis, results from Study 1 showed that processing load increases at the breakpoints between action units, and that processing load increases even more upon completion of several actions that can be integrated into a single unit. Increased processing at breakpoints might reflect something like “wrap-up” time—a phenomenon found in studies on discourse processing that is believed to reflect the progressive integration of units across time.

Studies 2 and 3 explored alternative reasons why processing load increases at breakpoints. Breakpoints correspond to relatively large amounts of change in the physical orientation and movements of an actor. These changes at breakpoints attract attention. Changes at breakpoints may also be informative, because they correspond to meaningful changes in both the actor’s body and in the objects in a scene. This could explain why people look longer at breakpoints even when they are presented out of sequence. In the present studies, however, increased processing load at breakpoints could not be fully explained by either physical changes or relative informativeness. It seems likely that on-line action understanding requires a complex interplay between perceiving changes, interpreting those changes, and integrating them into an ongoing representation of the actor’s goals and subgoals.

Acknowledgments

This research was funded by a grant from the *Norman H. Anderson Research Fund.* Thanks to Chris Pratt, Toti Moreland, and Allison Woo for help in data collection and to Gordon Bower, Barbara Tversky, and Sandra Lozano for comments and insights.

References

Abbott, V., Black, J. H., & Smith, E. E. (1985). The representation of scripts in memory. *Journal of Memory and Language, 24,* 179-199.

Baird, J. A., Baldwin, D. A., & Malle, B. F. (2006). *Parsing the behavior stream: Evidence for the psychological primacy of intention boundaries.* Manuscript under review, University of Oregon.

Baldwin, D. A., & Baird, J. A. (1999). Action analysis: A gateway to intentional inference. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life.* Mahwah, NJ: Lawrence Erlbaum Associates.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11,* 177-220.

Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology, 17,* 324-363.

Haberlandt, K., & Graesser, A. C. (1989). Processing of new arguments at clause boundaries. *Memory & Cognition, 17,* 186-193.

Haberlandt, K., Graesser, A. C., Schneider, N. J., & Kiely, J. (1986). Effects of task and new arguments on word reading times. *Journal of Memory and Language, 25,* 314-322.

Hard, B. M., Lozano, S. C., & Tversky, B. (2005). *Hierarchical encoding: Translating Perception into Action.* Manuscript under review, Stanford University.

Lozano, S. C., Hard, B. M., & Tversky, B. (2005). *Perspective-taking promotes action understanding and learning.* Manuscript under review, Stanford University.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology, 28,* 28-38.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology, 12,* 436-450.

Newtson, D., Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology, 35,* 847-862.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69,* 344-354.

Zacks, J., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130,* 29-58.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension.* New York: Academic Press.