Acta Psychologica 158 (2015) 8–18

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Simple visual cues of event boundaries

Tibor Tauzin *

Cognitive Development Center, Central European University, Hungary
Department of Cognitive Science, Budapest University of Technology and Economics, Hungary

ARTICLE INFO

Article history:
Received 28 August 2014
Received in revised form 14 March 2015
Accepted 23 March 2015
Available online 8 April 2015

JEL Classification:
2323
2340
2343

Keywords:
Event
Event primitive
Event segmentation
Memory

ABSTRACT

A stream of sensory information is organized into discrete temporal units through event segmentation. On the basis of several studies measuring participants' explicit decisions about event boundaries, some theorists suggest that this segmentation is induced by increased unpredictability. Since this approach cannot describe the segmentation of unfamiliar events, we assumed that event segmentation might be perceptually driven. We hypothesized that when a new event-relevant object is represented, it triggers event segmentation. In addition to explicit decisions, we measured memory performance, since it has previously been found to be a strong indicator of event segmentation. We presented simple videos to the participants in which geometric objects were flashing consecutively while an unpredictable change occurred. In the New Object condition flashing objects were replaced, while in the Same Object condition one non-kind-relevant feature of the objects was changed. In Experiment 1 the participants' task was to press a button when they detected a meaningful change in the stimuli. In line with the predictability-based theories, we found that both changes triggered the detection of an event boundary. To contrast our hypothesis with the predictions of earlier theories, in Experiments 2 and 3 memory accuracy was measured using the stimuli of Experiment 1. We only found a significant change in memory accuracy in the New Object condition, which suggests that the appearance of an event-relevant object can induce segmentation on its own, and indicates that the explicit-decisions methodology might lead to the improper conclusion that event segmentation is solely based on predictability.

"… to make an end is to make a beginning"
[T. S. Eliot, Little Gidding]

1. Introduction

1.1. Event segmentation

The process through which humans organize the continuous flux of sensory information into discrete units and through which they are able to recognize the beginning and the end of an event, has major relevance to perception both in- and outside the laboratory. In their experiments researchers often use separated events (trials and blocks) to provide multiple occurrences of the stimuli. These experimental designs are based on the assumptions that participants segment the experiment into smaller meaningful units that makes it possible to measure their performance after each unit. In everyday life, we also segment our experiences into episodes ranging from relatively huge time spans (such as listening to a talk) to extremely brief periods of time (e.g., recognizing a shooting star in the sky). In all such cases, the organization of continuous information into smaller meaningful units has a crucial role since the cognitive mechanism dedicated to this process enables us to exhibit different affective, cognitive or behavioral responses in correspondence with diverse events.

An event can be defined as a limited portion of space–time (Quine, 1985/1996). In psychological terms the perception of a Quine-event (or simply: of an event) was usually defined on the phenomenological level as 'a segment of time at a given location that is conceived by an observer to have a beginning and an end' (Zacks & Tversky, 2001) without constraints on their content (Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Although the phenomenological description of this process raises some questions (see later), event segmentation is claimed to be the mechanism that divides the continuous perceptual information into meaningful units (Kurby & Zacks, 2008), where the end of a given event necessarily implies the beginning of a new one. Hereafter, the characteristics of this process will be investigated in the visual domain regardless of empirical and theoretical contributions from the fields of auditory event segmentation and narrative understanding.

Participants in visual event parsing experiments were able to mark event boundaries (the so-called breakpoints) by pressing a button (for further details see: Kurby & Zacks, 2008; Newtson, 1973; Newtson,

* Cognitive Development Center, Central European University, Október 6. (6th of October) street 7, 1051 Budapest, Hungary. Tel.: +36 13273000x3663; fax: +36 18875010.
E-mail address: tauzint@ceu.hu.

http://dx.doi.org/10.1016/j.actpsy.2015.03.007
0001-6918/

Rindner, Miller, & LaCross, 1978) when a meaningful change occurred in the stimuli (Newtson, Engquist, & Bois, 1977). As these responses showed high reliability (Newtson et al., 1977; Speer, Swallow, & Zacks, 2003) the paradigm became widespread in the field.

1.2. Theories of event segmentation

Several theoretical approaches emerged during the past decades of research on event segmentation to describe what triggers event parsing.

According to the Event Segmentation Theory or simply EST (Kurby & Zacks, 2008; Zacks & Tversky, 2001; Zacks et al., 2007) event segmentation is an automatic (Noble et al., 2014; Sridharan, Levitin, Chafe, Berger, & Menon, 2007; see also Kurby & Zacks, 2008), multimodal control mechanism that regulates the allocation of cognitive resources to maintain or reset the so-called event model. The event model is the representation of the current event maintained in working memory. The purpose of an event model is to generate accurate perceptual predictions about perceived actions based on prior knowledge. While the predictions are accurate, the event model is the source of perceptual constancy. According to EST, on event boundaries the unpredictability increases resulting in a transient increase in prediction errors. For this reason an error detection mechanism resets the current event model and increases the influence of sensory input on perceptual processing in order to construct a new predictive event model in interaction with previously stored event schemata. The increased influence of sensory input results in the better encoding of event boundaries compared to non-boundaries. If an event model is not active any more, it will be moved to long-term memory. Thereafter, only the most thoroughly encoded parts of these event representations (its breakpoints) can be retrieved accurately from long-term memory.

The arguments of EST imply that, after numerous repetitions, a previously unpredictable change between events becomes more predictable and thus lets these previously separate events be unified in a single event schema. This prediction shows some overlap with a previous hypothesis about event demarcation, the so-called cut hypothesis (Avrahami & Kareev, 1994). Avrahami and Kareev (1994) suggested that a new event emerges every time a sequence of stimuli is repeatedly perceived in different contexts. Multiple occurrences of an event may increase predictability within an event and decrease predictability at the end of it.

The crucial role of predictability was also emphasized from a connectionist viewpoint. Hanson and Hanson (1996) provided evidence that similarity between the perceptual input and the currently maintained event model (the activated schema) together with the predicted specificity and duration determined whether an event boundary was perceived. Event segmentation was successfully modeled using one active schema and many stored ones, which is analogous to what EST hypothesizes. Furthermore, a more recent article from Reynolds, Zacks, and Braver (2007) directly tested the claims of EST from a connectionist point of view. They found that event segmentation can be modeled based on the mechanisms that EST proposes.

1.3. Problems with the theoretical approach of event segmentation

Theories of event segmentation converge on the assumption that prior knowledge and predictability are the two key factors that drive event segmentation. This implies that the temporal units of events that are similar to a given event schema are separated by transient time periods when predictability is low. Although the latter assumption seems plausible in most of the cases, it faces some major issues.

First, the level of predictability is unidentifiable in the case of events that are unfamiliar to the observer. For instance, during the synchronization of carburetors the different steps (in other terms the different events) could not be identified due to the lack of prior knowledge, however if the observer tracks which tool is used by the mechanic it could help to segment the event. Accordingly, it is known from a previous empirical study (Zacks, Tverky, & Iyer, 2001) that observers have no difficulty segmenting videos of unfamiliar actions. Thus, it is possible that tracking the objects involved in an event is sufficient for event demarcation.

Second, in the case of frequent events an event becomes more predictable due to repetition. Therefore, predictability based event segmentation theories should argue that the number of event boundaries (which are characterized by unpredictability according to them) will decrease in these events. Although one could also argue that perceptual predictions should not be affected by previous knowledge, Zacks et al. (2001, Experiments 2 & 3) showed that, after training, participants recognized significantly *more* fine-grained event boundaries in an event which contradicts both predictions (but see: Bläsing, 2015; Hard, Tversky, & Lang, 2006).

Third, empirical evidence suggests that unpredictability might not always be enough to trigger segmentation even if the context is familiar to the observer, thus prior knowledge is available to determine what is unpredictable. Carroll and Bever (1976) showed that the recognition speed of film excerpts followed by an unpredictable change in action was slower than that of a film excerpt followed by an unpredictable change in the viewpoint of the camera. This result suggests that only the former change had an effect on memory which can be an indirect indicator of event parsing (see below). In a later study Schwan, Garsoffky, and Hesse (2000) found a similar effect: an unpredictable cut did not change memory accuracy. This suggests that a basically unpredictable change (a film cut) which is atypical in natural contexts, thus could not be the part of stored event schemata, is not sufficient to trigger event parsing.

The above problems indicate that changes in the level of predictability cannot account for segmentation even if prior knowledge is available. Unpredictability might co-occur with breakpoints but does not necessarily indicate it, which implies that simple visual cues can be more crucial in event segmentation than it was previously suggested.

1.4. Measuring visual event segmentation through its effect on memory performance

In the course of the behavioral paradigm, that has been extensively used to measure event segmentation since Newtson's (1973) pioneering study, participants need to press a button 'when a (meaningful) unit ends and a different one begins' in the stimuli. Although the vast majority of subsequent studies (e.g., Ginsburg & Smith, 1993; Hanson & Hirst, 1989; Hard et al., 2006; Lassiter, Geers, & Apple, 2002; Massad, Hubbard, & Newtson, 1979; Newtson & Rindner, 1979; Newtson et al., 1978; Schwan & Garsoffky, 2004; Schwan et al., 2000; Zacks, 2004) used this direct measure to determine where event boundaries are perceived, this paradigm is not suitable for disentangling the detection of a conspicuous change in the stimuli and the recognition of the start of a new event. For instance, an unexpected change in the location or position of an object can be conspicuous, and in some cases, it can indicate a breakpoint. However, the change in location or position is not meaningful in most cases and it presumably does not trigger event parsing.

It is theoretically possible that the indirect measurement of event segmentation could differentiate the conspicuous change and event parsing. Previous studies raised the possibility that memory accuracy could be an indirect indicator of event parsing. In an early experiment Newtson and Engquist (1976, Experiment 3) revealed that after participants watched short films, recognition accuracy was significantly higher for frames that appeared on event boundaries than frames that did not. This effect was also found in children (Newtson, 1998). The finding that the details of an event on its boundaries are retrieved better suggests that these periods were encoded better, presumably due to their higher salience. Further research showed that besides recognition (Lassiter, 1988; Lassiter, Stone, & Rogers, 1988) recall performance was also better on event boundaries (Schwan & Garsoffky, 2004). In accordance with the former findings Swallow, Zacks, and Abrams