



Segmentation in Cinema Perception

Author(s): J. M. Carroll and T. G. Bever

Source: *Science*, New Series, Vol. 191, No. 4231 (Mar. 12, 1976), pp. 1053-1055

Published by: [American Association for the Advancement of Science](#)

Stable URL: <http://www.jstor.org/stable/1741429>

Accessed: 16/02/2011 11:19

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=aaas>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Association for the Advancement of Science is collaborating with JSTOR to digitize, preserve and extend access to *Science*.

<http://www.jstor.org>

the egg mass during isoproterenol infusion indicates that the cilia provide the sole propellant force under these experimental conditions. Indeed, the cilia by themselves appear to be capable of transporting the ovum to the site of fertilization within normal time limits. Obviously the rapid muscle-induced movements are not necessary for transport.

There are at least three ways of interpreting the observations as to the mechanisms underlying ampullary egg transport. (i) The ciliary and muscular influences could be independent and their propellant forces additive. The rapid muscle-induced movements would therefore be randomly superimposed on slower ciliary transport and ineffective for net transport. (ii) At the opposite extreme, the forces exerted by the cilia might act on the egg mass only when muscle forces are negligible, that is, the cilia-induced and muscle-induced movements are mutually exclusive. In such case, the cilia may completely compensate for the absence of muscular activity if they are as effective as the muscle in transporting ova. (iii) There may be some intermediate degree of interaction between muscular activity and the ciliary apparatus. The roles of both muscle and cilia might be variable and dependent on various properties of the oviductal wall, luminal surface, and intraluminal fluid and, thus, be controlled or regulated by the sex hormones.

To explore further the third interpretation, we studied the interaction of the muscle and cilia under experimental hormonal control. We chose a hormonal state associated with remarkably enhanced muscular activity, dramatic rapid movements of the cumulus egg masses, and rapid ampullary transport (8). Castrate rabbits were treated with estrogen for 4 days; for the 24 hours before the experiment, these estrogen-primed subjects were also treated with CI628, a nonsteroidal estrogen antagonist. If the exceptionally rapid rate of transport induced by hormonal treatment were due to alteration of the muscle activity, dramatic slowing of transport would follow pharmacological inhibition of the muscle.

In order to characterize transport through the ampulla as precisely as possible, we analyzed 16-mm motion pictures of the CI628 experiment and, with the help of a computer, plotted a graph of the movement of the egg through the ampulla (Fig. 1). Transport under the influence of isoproterenol was, as expected, relatively free of high velocity movements. Totally against expectations, the transport was of shorter total duration. The average velocity is about 0.25 mm/sec, substantially greater than the velocities measured under the influence of isoproterenol in the experi-

ments on estrous and ovulatory rabbits.

Although these data also support the hypothesis that muscle contractions per se do not contribute to net ampullary transport, they do not preclude a functional role for the muscle; the motion imparted by the muscle to the egg may play some other important role in fertility. The alteration of ciliary transport by hormonal manipulation, however, is clearly apparent. These studies emphasize the importance and relative autonomy of the ciliary apparatus in effecting ovum progress within the mammalian oviduct. Further studies aimed at selective control appear desirable.

S. A. HALBERT

P. Y. TAM*

R. J. BLANDAU

Department of Biological Structure,
University of Washington, Seattle 98195

References and Notes

1. R. J. Blandau, in *Handbook of Physiology*, Section 7: *Endocrinology*, R. O. Greep and E. B. Astwood, Eds. (Williams & Wilkins, Baltimore, 1973), vol. 2, part 2, pp. 153-163.
2. M. J. K. Harper, *Proc. Int. Congr. Anim. Reprod.* 4, 375 (1961); *J. Reprod. Fertil.* 2, 522 (1961).
3. R. J. Blandau, in *The Mammalian Oviduct*, E. S. E. Hafez and R. J. Blandau, Eds. (Univ. of Chicago Press, Chicago, 1969), pp. 129-162.
4. ———, in *Methods in Mammalian Embryology*, J. C. Daniel, Jr., Ed. (Freeman, San Francisco, 1971), pp. 1-14.
5. S. A. Halbert, P. Y. Tam, R. J. Adams, R. J. Blandau, *Gynecol. Invest.*, in press.
6. G. L. Duff, H. F. Stegall, C. Nelson, R. Blandau, J. Boling, *Proc. Conf. Eng. Med. Biol.* 14, 198 (1973); S. A. Halbert, J. L. Boling, R. J. Blandau, J. A. Ringo, paper presented at the 28th Annual Conference on Engineering in Medicine and Biology, New Orleans, Louisiana, 20 to 24 September 1975.
7. J. L. Boling and R. J. Blandau, *Biol. Reprod.* 4, 174 (1971).
8. R. J. Blandau and J. L. Boling, unpublished data.
9. Supported by NIH contract No. HD-2788.
- * Present address: Department of Mechanical Engineering, University of California, Berkeley 94720.

25 September 1975; revised 24 December 1975

Segmentation in Cinema Perception

Abstract. *Viewers perceptually segment moving picture sequences into their cinematically defined units: excerpts that follow short film sequences are recognized faster when the excerpt originally came after a structural cinematic break (a cut or change in the action) than when it originally came before the break.*

A central theme in the study of the cinema interprets the structure of film as a metaphorical "language" (1). Such a conceptualization has apparent validity. A motion picture sequence attempts to portray an event by imposing a structure of cuts, zooms, tracks, pans, framings, and the like on that event. Analogously, a sentence represents an idea by imposing a syntactic and phonological structure on a set of lexical items. The task of the viewer of a film sequence is to apprehend the event represented in the sequence. The analogous task of the listener is to recover the idea encoded in the sentence.

Up to now, the metaphor of film as language has been restricted to purely theoretical accounts of cinema. We show that the methodologies used in the psychological investigations of sentence perception can give experimental support to this characterization. Our report demonstrates that just as the syntactic structure of sentences plays an organizing role in their segmenta-

tion into major processing units (2), the structural organization of a film sequence plays an analogous part in segmentation of films.

We examined two aspects of organization in motion picture sequences, events and scenes. Events are sequences of constituent actions (3). Scenes are sequences of cinematic shots (4).

We showed viewers short film sequences, each of which was divided into two segments by a change in the action (type E), by a cut (type C), or by both (type EC). After the film sequence we presented the viewers with brief excerpts (probes) from either the first or second segment of the sequence. Viewers recognized probes from the second segments faster than those from the first, indicating that the material in the first segment is organized separately and more abstractly than the material in the second segment (5). This demonstrates that cinematic structure organizes viewers' perceptual processing. Furthermore, it demonstrates the validity of applying psycholinguistic techniques to the study of an art form.

The film sequences depicted ordinary daily experiences, such as conversation, scenes of people walking around in rooms or outdoors, traffic scenes, or meals. The cuts in the type EC and the type C sequences involved changes in camera distance (medium shot to close shot or medium shot to long shot) or a change in camera angle (a lateral change of 45° or 90°)

Table 1. Mean affirmative decision times.

Type of stimulus sequence	Source of probe	
	First segment (msec)	Second segment (msec)
EC	951	816
C	906	895
E	1055	917
Total	971	876

View fixation field	View stimulus film sequence	Red fixation field	View probe excerpt	Indicate if probe was part of stimulus sequence	Categorize stimulus sequence as old, new, or paraphrase
---------------------	-----------------------------	--------------------	--------------------	---	---

Fig. 1. Experimental procedure for each trial. Each subject viewed 30 such trials.

(6). Four scenes of each type were constructed. In two of these, the probe was the six frames just before the structural break, and in the other two, it was the six frames after the structural break.

Thirty volunteers from psychology classes at Columbia University served as viewers. The presentation of every stimulus item was preceded by the display of a gray field with a central black dot that was intended to serve as a fixation field (Fig. 1). The test sequence was then displayed, followed by a six-frame presentation of the fixation dot on a red field (approximately 375 msec). The red field was the only color in the stimuli and served both to mark the end of the stimulus sequence and to ready the viewer for the probe display (7).

Following the fixation field-film sequence-fixation field presentation a probe excerpt of six frames (375 msec) was displayed. If the six-frame probe sequence originated in the directly preceding film sequence the viewer was to respond "yes." If the probe was not an excerpt from the film sequence, the viewer was to say "no." A photocell relay circuit started a millisecond clock at the onset of the probe. The viewer's vocal response stopped the clock. The measured time interval was taken as the decision time.

After six practice trials (a set of three trials repeated), each viewer was shown a series of 27 test trials in one of two complementary presentation orders. In 12 of the 27, the probe was an excerpt from the stimulus sequence, and in the remainder it was not. Each subject was informed that the major focus of the experiment was the decision time. However, as an ancillary task, viewers were next asked to classify the scene they had viewed as old, new, or a paraphrase. An old scene was a literal repetition of a scene previously displayed in the test series; a new scene was one not yet presented; a paraphrase scene was defined as one that was technically new but similar to an old one. The purpose of this second task was to ensure that viewers gave full attention to processing the film sequences. Mistaking old for new scenes (or vice versa) was rare.

Since presentation order had no reliable effect, we pooled the data for analysis. Our main prediction, that decision time for second-segment probes is shorter than that for first-segment probes, was confirmed (Table 1) (8).

A second result is that both cuts and changes in the action define effective segmentation boundaries. The main prediction was confirmed when we separately consider those sequences having event boundaries (types E and EC) and those having cuts (types C and EC) (9). Thus the segmentation unit in film processing is determined by an interaction of (at least) two levels of film structure.

Finally, we found that decision time for second-segment probes is significantly shorter than it is for first-segment probes in both type EC and type E cases (10), but not for type C comparisons. The differences between the results for types E and EC scenes and type C scenes suggest an elaboration of our structural model of film processing: cuts which coincide with event boundaries stimulate segmentation in processing, whereas those that do not change the event do not stimulate segmentation.

Although this pattern of results supports our notion of film "language" we must also consider other interpretations. The possibility that viewers verify second-segment probes faster simply because those probes appear in the sequence closer to presentation of the probe cannot provide a comprehensive account of the data, since we found no difference between first and second segments for type C sequences. In addition, serial position cannot account for the data for type E and type EC sequences, since the actual test sequences with first-segment probes also have slightly longer first segments than the sequences with second-segment probes. The second-segment probes actually originate slightly earlier in the scene sequences than first-segment probes (11).

This fact, however, suggests another interpretation of the results. Suppose that serial position is the effective variable but that the further the original occurrence of the probe is from the probe display, the faster the decision time. Presumably, this claim would depend on the inferred (memory) interference of the intervening material. We correlated the mean decision time across subjects with the number of frames from the end of the original probe site in the test sequence to the end of the test sequence itself. This correlation is in the right direction but it is not statistically significant (12). Indeed, that it is even weakly in the right direction appears to be due only to our main effect.

When the same correlation is performed separately for the second-sequence probe cases and first-sequence probe cases we obtain correlations in the opposite direction (13).

If the effective variable is actually the position of the probe site with respect to the break, the cut might cause backward masking, which perceptually partially obliterates material immediately preceding the cut in the scene sequence. This is reasonable given that abrupt visual discontinuities like cuts disrupt normal eye movement (14). According to this explanation, the first-segment probes take longer to verify because they have been partially obliterated by backward masking and not because they originate in the initial, and therefore less accessible, segment. This could explain the results for type EC scenes, but it is not consistent with the results for type C or type E scenes. For type E scenes it would predict no difference between first-segment and second-segment latencies because there is no cut; however, the decision time difference is significant. Conversely, the backward masking explanations predicts a first-segment versus second-segment probe latency effect for type C scenes; however, there is no significant latency difference.

We conclude that the structural organization of film sequences can play an organizing role in their segmentation. In particular, we have initially confirmed that the structural units of film sequences are the segmentation units of film processing.

The tentative model for the perceptual processing of film suggested by this research is similar to the psycholinguistic model of sentence perception (15). Viewers initially process material in real time until they reach a structural boundary. At that point all material is recoded into a more abstract holistic form and dismissed from immediate memory (that is, viewers perceptually segment the sequence). Viewers then proceed to the next unit.

A more complete model of film processing must await further research. However, these results encourage us in the view that the experimental and structural methods of modern psycholinguistics can be usefully applied to the psychological study of cinema.

J. M. CARROLL
T. G. BEVER

*Department of Psychology,
Columbia University, New York 10027*

References and Notes

1. S. Eisenstein, *Film Form* (Harcourt, Brace & World, New York, 1949); B. Balazs, *Theory of the Film* (Dover, New York, 1970); A. Bazin, *What Is Cinema?* (Univ. of California Press, Berkeley, 1971).
2. J. Fodor and T. Bever, *J. Verbal Learn. Verbal Behav.*

- hav. 4, 414 (1965); T. Bever, J. Lackner, R. Kirk, *Percept. Psychophys.* 5, 4 (1969).
3. R. Barker and H. Wright, *The Midwest and Its Children* (Harper & Row, New York, 1955); H. Dickman, in *The Stream of Behavior*, R. Barker, Ed. (Appleton-Century-Crofts, New York, 1963), pp. 23-41; C. Metz, *Film Language: A Semiotics of the Cinema* (Oxford Univ. Press, New York, 1974); J. Carroll, in *Studies in Structural Aesthetics*, T. Bever, Ed. (Erlbaum, Hillsdale, N.J., in press).
 4. K. Reisz and G. Millar, *The Technique of Film Editing* (Hastings House, New York, 1968).
 5. This is analogous to the fact that a probe word from the second clause of a sentence is recognized faster than one from the first clause of the sentence [D. Caplan, *Percept. Psychophys.* 12, 73 (1972); J. Kornfeld, thesis, Massachusetts Institute of Technology (1973); L. Shedletsky, thesis, University of Illinois (1974); see also G. Suci, P. Ammon, P. Gamlin [*Lang. Speech* 10, 69 (1967)].
 6. The segments of the type E sequences were linked by mise-en-scène devices other than cuts, such as, pans, tilts, dollies. The devices linking the segments were designed not to be phenomenally obtrusive (for example, so-called jump cuts or expressive montage) nor were they invisible in the sense of Hollywood "hidden" cuts. Subjects did not report afterwards that their attention had been drawn to the cuts and other devices, but many were aware that there had been cuts in the sequences. The devices seemed to be similar with respect to phenomenal obtrusiveness, but we have no way of precisely quantifying their relative strengths.
 7. Black-and-white motion picture sequences were rear projected by an L & W Data Analyzer (model 224A). The film sequences were projected at 16 frames per second. The film sequences ranged from 75 frames to 100 frames in length (from 4.7 to 6.3 seconds in duration). Viewers watched the back-projection screen at a distance of 3.5 feet. The image was 10.5 inches by 8.5 inches; it subtended a visual angle of 11.5° by 14°.
 8. By subject, $P < .003$ (one-tailed, Wilcoxon matched pairs signed ranks) and by item, $P < .05$ (one-tailed, Wilcoxon rank sums).
 9. By subject, $P < .01$ and $P < .02$ (one-tailed). We have too few items in each subcategory to perform statistics by item in the internal analyses.
 10. By subject, $P < .01$ and $P < .01$ (one-tailed).
 11. The mean distance from the end of the probe site to the end of the test sequence in the second-segment probe items is 49 frames. For first-segment probe items the mean distance is 29 frames (a difference of 1.25 seconds).
 12. Across subjects, $r = -.12$, $t = 0.37$, d.f. = 10.
 13. Second-sequence probe cases, across subjects, $r = +.82$; first-sequence probe cases, across subjects, $r = +.51$.
 14. M. Potter and E. Levy, *J. Exp. Psychol.* 81, 10 (1969).
 15. J. Fodor, T. Bever, M. Garrett, *The Psychology of Language* (McGraw-Hill, New York, 1974); J. Carroll and T. Bever, in *Handbook of Perception*, E. Carterette and J. Friedman, Eds. (Academic Press, New York, in press), vol. 7.
 16. We gratefully acknowledge advice we received from J. Hochberg and L. Gellman.

9 June 1975; revised 24 October 1975

Brain Wave Components of the Contingent Negative Variation in Humans

Abstract. *In a contingent negative variation paradigm with two stimuli paired at an interstimulus interval of 4 seconds, two distinct waveforms having functional and topographic differences are observed. An early wave is maximal over the frontal cortex and is elicited by the warning stimulus. A later wave, maximal over the motor cortex, precedes the imperative stimulus and is identified with preparation for motor response.*

Slow brain potential shifts are recorded from the human scalp during periods of preparation or active mental involvement. The most familiar of these potentials, the contingent negative variation (CNV), first described by Walter *et al.* (1), occurs during the interval between a warning stimulus and a subsequent imperative stimulus that requires a mental or motor response. The CNV has proven sensitive to a number of psychological variables, none of which can yet be singled out as a definitive or exclusive correlate (2). This lack of consensus as to the underlying variables reflects a growing belief that the CNV may comprise two or more separate processes, combining in unrecognized proportion (3-5). Here we distinguish two component processes on the basis of latency, topography, and functional properties. An early component is maximal over the frontal cortex and appears in response to the warning stimulus—whether or not it is paired with a subsequent imperative stimulus. A second, later component shares critical features with the readiness potential, first described by Kornhuber and Deecke (6) as a negative shift preceding voluntary movement, in that it has a more central distribution over motor areas of the cortex and is related to the dynamics of the motor response.

This research led from our previous findings that the distribution of the CNV over the scalp varied in accordance with different perceptual and motor tasks (7). In one condition, in which the word "left" or "right" served as a warning stimulus and designated the hand to be used in the response, the bilaterally recorded CNV became progressively asymmetrical throughout the 1.5-second interstimulus interval, being greater over the contralateral hemisphere. The increasing asymmetry before the signal to respond suggested some contribution from the readiness potential, which has been observed to lateralize in a similar manner and over a similar time course (8). However, this contribution has been difficult to assess, especially in view of reports that CNV characteristics sometimes have little to do with specific response requirements (9).

A means whereby the readiness potential might be observed separately from other negative waves was suggested to us by recent reports by Loveless and Sanford (4, 5) and Weerts and Lang (10), who have shown that the CNV assumes two distinct and nonoverlapping phases when the interstimulus interval is extended beyond the typical range of 1 or 2 seconds. The first phase is a broad negative wave that peaks

within a second or so after the warning stimulus. Under conditions in which the interstimulus interval is constant and responding speed is emphasized, a second negative shift appears before the imperative stimulus.

We were able to observe clearly both these waves in a reaction time task employing tone bursts paired with flashes, separated by a relatively long interstimulus interval of 4 seconds. Subjects were required to respond to the light flashes as quickly as possible, but not prematurely, by fully depressing a telegraph key beneath the index finger. Complete depression of the key (3-mm throw) required a force equivalent to 450g; however, reaction times were measured from the earliest key displacement. Identical keys were located under each hand, and the appropriate hand on each trial was designated both by the pitch of the warning tone, which was either 1000 or 1500 hertz, and by the shape of the corresponding flash, which was either a square or a diamond (square rotated 45°). The paired stimuli were presented in random order. Subjects reported that the two pitches were easily discriminable and enabled unimanual response preparation during the interstimulus interval. The different combinations of high or low tone pitch, square or diamond flash shape, and left or right responding hand were balanced over subjects.

The tone bursts were rectangularly gated for a duration of 20 msec and delivered over binaural insert earphones at an intensity of 85 db. The flash patterns were formed by masking the milk-glass faces of two Grass PS-2 photostimulators, mounted in different channels of a three-field tachistoscope. The third channel was dark except for a dim fixation cross. At a viewing distance of 175 cm, the square and diamond each subtended a visual angle of 2.3° by 2.3°. The original intensity, at a setting of 4, was reduced approximately 80 percent through the optics of the tachistoscope.

Fourteen right-handed psychology students (nine male) served as subjects. During a single 2-hour session they received two identical blocks of 80 paired tones and flashes, half requiring a right-hand response and half a left-hand response. The interval between the key press and the warning tone for the next trial varied randomly from 3.5 to 7.5 seconds.

The blocks of paired stimuli were preceded by a condition in which subjects listened attentively but without responding to 30 unpaired tone bursts, varying randomly between the high and low pitches and separated by random intervals of 3.5 to 7.5 seconds. In an additional condition interposed between the two blocks of paired stimuli,