



# LUND UNIVERSITY

## The objects of attention: Causes and targets

Brinck, Ingar

*Published in:*  
Behavioral and Brain Sciences

*DOI:*  
[10.1017/S0140525X03270072](https://doi.org/10.1017/S0140525X03270072)

2003

[Link to publication](#)

*Citation for published version (APA):*

Brinck, I. (2003). The objects of attention: Causes and targets. *Behavioral and Brain Sciences*, 26(3), 287. <https://doi.org/10.1017/S0140525X03270072>

*Total number of authors:*  
1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

is an empirical question. Hopefully, someone will be able to come up with ways to test them empirically. Until then, none is more than a hopeful hypothesis.

## Grammar originates in action planning, not in cognitive and sensorimotor visual systems

Bruce Bridgeman

Department of Psychology, University of California, Santa Cruz, CA 95064.  
bruceb@ucsc.edu <http://psych.ucsc.edu/Faculty/bBridge.shtml>

**Abstract:** While the *PREDICATE(x)* structure requires close coordination of subject and predicate, both represented in consciousness, the cognitive (ventral), and sensorimotor (dorsal) pathways operate in parallel. Sensorimotor information is unconscious and can contradict cognitive spatial information. A more likely origin of linguistic grammar lies in the mammalian action planning process. Neurological machinery evolved for planning of action sequences becomes applied to planning communicatory sequences.

It is tempting to relate ideas in linguistics with ideas in neurophysiology, because at base much of linguistics is about the design and operation of a neurophysiological machine, in the language areas of the brain. In the spirit of consilience, such efforts are necessary. Hurford's effort at tying together a formal logic developed within linguistics with the interactions between brain areas echoes another effort in *BBS* a decade ago (Landau & Jackendoff 1993). Like that effort, though, this one founders by proposing a parallel that on closer examination turns out to be illusory.

The heart of Hurford's article, identifying the dorsal and ventral streams of visual processing with a logical *PREDICATE(x)* structure, misses the mark because there is a tight logical relationship between subject and predicate; but information in the two visual streams can be independent and even contradictory, running in parallel to subservise different functions (Bridgeman et al. 2000; Milner & Goodale 1995). The dorsal/ventral terminology is somewhat deceptive, for some cortical areas that are anatomically dorsal to the primary visual cortex are shared by both pathways, or even belong to the "ventral" pathway. Terms that capture the contrast in the functions of the two pathways are more useful. *What* and *where*, as used by Hurford, are misleading because both pathways carry useful where information – it's just that the where information in the two pathways is sometimes contradictory. The terms *cognitive* and *sensorimotor* are preferable, as they describe the distinct functions of the pathways, whereas Milner and Goodale (1995) suggest *what* and *how*. Functional terms are preferable because they are less likely to lead to oversimplification or overinterpretation.

The idea of using logical grammar notations developed within linguistics for describing information processing in the brain is a productive one that promises to enrich neuroscience. The literal application of logical structures to describe information processing within the two-visual-systems context, however, is wide of the mark because the linguistic structures and the logical structures of visual architecture are not parallel.

Subject and predicate are both conscious in the minimal sense that one can talk about them. Their identities and relationship can be described, their application can be planned in language, and they define inseparable parts of a single linguistic act. The sensorimotor pathway, however, can function without cognitive participation and without conscious intervention all the way from stimulus to response, an example of "vertical modularity" (Bridgeman 1999).

A recent method of dissociating cognitive and sensorimotor information exploits the Roelofs effect (Bridgeman et al. 2000), without confounds from motion of the eye or of the visual stimulus. The Roelofs effect is a tendency to misperceive the position of a target presented along with an off-center background. A rec-

tangular frame offset to an observer's left, for example, causes the position of a target presented within the frame to be mislocalized to the right. Despite this mislocalization, observers could jab the target accurately, without the frame affecting their behavior. The effects may be due to the frame biasing the observers' subjective straight ahead, stored unconsciously in a sensorimotor system.

Anatomical connections between dorsal and ventral streams do not contradict the separability of their functions, any more than communication between two people contradicts their distinctness. Communication between the two streams is needed to initiate action (usually a cognitive-system function), to monitor progress in the execution of the action, and to modify goals of actions.

Rizzolatti and Arbib (1998) also use the language of language to describe neurophysiological relationships, but they explicitly specify a prelinguistic grammar to distinguish it from spoken language. Thus their "grammar" refers only to a set of rules by which the brain processes information. It is unrelated to language in the usual sense. Semantics in vision and in neurophysiology refer to a relation of images with meaning; its relation to language is more metaphorical than literal.

If the distinction between cognitive and sensorimotor pathways of the visual system does not offer a source for the evolution of the logical relations necessary for language, what does? A more likely alternative is the planning process that all mammals possess and that becomes particularly important and well-developed in primates. Plans for action exist separate from the sensory or motor worlds, and their steps must be executed in a particular order to be effective. Grammar may have appropriated an existing capability for planning of action sequences to the planning of communicatory sequences (Bridgeman 1992). Language, then, is a new capability built mostly of old parts, but the parts originate in motor planning, not in visual coding.

## The objects of attention: Causes and targets

Ingar Brinck

Department of Philosophy, Kungshuset Lundagård, Lund University, SE-222 22 Lund, Sweden. [ingar.brinck@fil.lu.se](mailto:ingar.brinck@fil.lu.se)  
<http://213.80.36.53/temp/filosofen3/staff/person.asp?id=17&lang=eng>

**Abstract:** The objects of attention can be located anywhere along the causal link from the source of stimuli to the final output of the vision system. As causes, they attract and control attention, and as products, they constitute targets of analysis and explicit comments. Stimulus-driven indexing creates pointers that support fast and frugal cognition.

Hurford suggests that the objects of attention should be understood as indexed, arbitrary objects identified by their location in a mental, spatial map. Objects of attention are available to the subject without categorisation or encoding of their properties or locations.

I do not agree with Hurford's characterisation of indexed objects as arbitrary and identified by their location in a mental map. First, indexing is not really arbitrary but is stimulus-driven. Not any object will be indexed, but only those that are salient enough to impinge on the subject. Indexing is caused by some property of the object, although that property will not be encoded (Pylyshyn 1999; 2000). Furthermore, at the moment of indexing, the objects are distinguishable as visual patterns or clusters in the visual field.

Finally, the spatial map is not mental, but the scene in the real world forms a local map that contains the indexed objects. The scene itself does not have to be memorized. Indexed objects serve as pointers that allow the subject to access and revisit locations in a distal environment without engaging attention. Thus, indexed objects support fast and frugal cognition, which exploits information in the environment (Brooks 1991; Hutchins 1995).

It is difficult to see how indexed objects could be objects of at-

tention. We can think of objects of attention as either causes attracting attention, or effects, that is, products of focal attention. Objects of attention conceived of as effects constitute enduring targets to which attention can be maintained. They are analysed, and if kept in short-term memory long enough to reach consciousness, they can be intentionally commented on by the subject (Weiskrantz 1997).

To allow for sustained attention, a target object must at least make it possible for the subject to track the object by its properties. The subject encodes the properties in short-term memory. Objects of attention cannot be discriminated by mere location, because the identification of locations relies on a previous segmentation of space (Driver & Baylis 1998). Minimally, target objects are constituted by segmented regions that form unities also when in motion. They have some spatiotemporal consistency.

I do not specifically address the question whether objects of attention are objects, features, or locations, but take it that attention is directed to objects (O'Craven et al. 1999; Yantis 1998). The objects of attention can in principle be located anywhere along the causal link from the source of the stimulus to the final output of the vision system. Which properties will be ascribed to the objects of attention depends on the level of analysis (Eilan 1998). The properties reflect the various cognitive roles of the objects of attention.

Objects of attention can be introduced on at least three levels of analysis. On an initial, preattentive level, they constitute the input to the early vision system and are best thought of as causes. This level processes segmented objects for focal attention and subsequent analysis. On a computational and attentive level, further on in the early vision system, objects of attention constitute targets that are processed in the dorsal and ventral paths. The dorsal and ventral paths may construct different and incompatible representations from stimuli from the same source, without the subject's noticing it (Goodale et al. 1994).

On a psychological, or phenomenological, level, which receives output from the early vision system, the objects of attention will be multimodal, three-dimensional percepts. Percepts occur on a personal level and are directly available for the organism as a whole, as opposed to being processed subpersonally. The subject may become consciously aware of them and choose to comment on them (Weiskrantz 1997). Comments are voluntary and intentional and can be communicated through behaviour or language. A comment will be cognitively penetrable if sensitive in a rational, or semantically coherent, way to the organism's goals and beliefs (Pylyshyn 1999).

Hurford furthermore suggests that the objects that subsequently are indexed attract attention, treating them as causes of focal attention. He claims that certain "natural attention-drawing properties" of the objects attract attention. These properties concern the biological needs of the subject and are highly encapsulated. In contrast to the percepts that are arrived at after an analysis in the ventral stream, these properties are not accessible to the subject on a personal level. Information about them is exchanged only between the subsystems of vision.

I do not see the need to introduce "natural attention-drawing properties" to account for attention attraction. I agree that whatever it is that attracts attention, it must be of interest to the subject. An "abstract" object cannot be so, simply because it is propertyless. Attention is attracted by objects that have an informational, and not merely causal, impact on the subject (Brinck 2001). They are at odds with what the subject is expecting on the basis of previous experience. But the impact is not necessarily related to biological needs, or to positive or negative values. Except for cases when an anticipated object attracts attention, the object will only receive a value for the subject once it has been detected (Eimer et al. 1996).

I submit that so-called goal-driven attention works top-down, in anticipation of some well-defined item. The subject is searching for a particular object, and the attention is geared to react when it appears (Ballard et al. 1997; Yantis 1998). The subject's needs and

desires determine the aim of the search. The salient feature that serves to indicate the appearance of the object is likewise selected before the search begins.

Stimulus-driven attention, however, works bottom-up. Attention is attracted by sudden and unexpected changes in the subject's immediate environment (Freyd 1987). Expectancy relates to familiarity. The change must introduce a new and somehow anomalous object or feature in the visual field to draw the attention of the subject (Yantis 1998; Yantis & Johnson 1990). It seems as well that the saliency of the object will increase if the object is behaviorally relevant according to the needs or drives of the subject (Gottlieb & Goldberg 1998).

To sum up, indexed objects as described by Hurford can only serve as pointers. If conceived of as objects with properties (albeit not encoded), they also take on the role of a cause that controls the subject by inducing her to index them. However, indexed objects can never be targets of attention. They are mere placeholders.

## What proper names, and their absence, do not demonstrate

Andrew Carstairs-McCarthy

*Department of Linguistics, University of Canterbury, Christchurch, New Zealand. [andrew.carstairs-mccarthy@canterbury.ac.nz](mailto:andrew.carstairs-mccarthy@canterbury.ac.nz)  
<http://www.ling.canterbury.ac.nz/>*

**Abstract:** Hurford claims that empty variables antedated proper names in linguistic (not merely logical) predicate-argument structure, and this had an effect on visual perception. But his evidence, drawn from proper names and the supposed inability of nonhumans to recognise individual conspecifics, is weak. So visual perception seems less relevant to the evolution of grammar than Hurford thinks.

Hurford draws attention to a parallel between, on the one hand, the roles of the ventral and dorsal pathways in vision, and, on the other, the roles of predicates and variables in predicate calculus. Just as the variable in predicate calculus has no role other than a deictic or indexical one, of locating an individual to which certain predicates belong, so the dorsal pathway (it seems) has scarcely any role other than to locate an object in space, nearly all its other characteristics being processed via the ventral pathway. How significant is this for language, either today or at an earlier evolutionary stage? Hurford does not claim that the correlation is today very close, and I agree with him. One cannot identify the dorsal-ventral contrast with the noun-verb contrast, for example. But he alleges a reflection of the dorsal-ventral contrast in the mental representations of all animals except modern humans, inasmuch as (he claims) only modern humans have a concept of individuals that are in principle proper-nameable – that is, individuals associated with more semantic content than mere indexical place-holders. On that, I find what he says unpersuasive. So I suspect that the parallel that he adduces has even less significance for language than he suggests. If so, then visual perception sheds little or no light, unfortunately, on the puzzle of why language (particularly syntax) is as it is.

"Protothought had no equivalent of proper names," says Hurford (sect. 1.3), and that is why it is easy to fool tern chicks about their parents: visually they are so easily fooled that they will react towards a loudspeaker as if it were a parent tern. Hurford concludes from this that tern chicks have no mental representation of their parents as individuals. (Hurford would presumably interpret in the same way the apparently sophisticated social awareness of vervet monkeys; see Cheney & Seyfarth 1990.) But that seems an overambitious conclusion. Terns may be easier to trick than humans are, but that proves nothing relevant to this issue. Let us suppose that, unbeknownst to me, Jim Hurford has an identical twin brother, Tim Hurford. I know Jim Hurford slightly from occa-