LUND UNIVERSITY

The objects of attention: Causes and targets

Brinck, Ingar

Published in:
Behavioral and Brain Sciences

DOI:
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

with the neural processing of event perception and motion perception, which I now take up.

Anderson & Oates join with **Werning** in rejecting my claim that the origins of perceptual predicates lie solely in the ventral stream. Anderson & Oates suggest that the dorsal pathway could produce representations to underlie predicates like *REACHABLE(x),* and the ventral pathway could produce representations to underlie predicates like *RED(x).* This has the same disadvantage as Werning's proposal, noted at the end of section R5, namely, that it does not provide any explanation for why the information coming through these separate streams should have the same predicate-argument format. The blindsight patient mentioned by Anderson & Oates could indeed reach accurately, a feat accomplished by his dorsal stream, but the property of reachability never got transferred upstream to mechanisms involved in reporting on events. The central claim of the target article is that only properties delivered by the ventral stream provide the predicates used in representations which, through memory, can become the basis for linguistic representations.

Werning argues that properties in the general super-category of *motion* are detected by the dorsal stream. Note first that there is, as **Woll**'s commentary mentions, little evidence that dorsal stream parietal systems are activated in sentence processing, even when space is referred to in spoken language. Thus, if we envisage the diagram in Figure R1 as a kind of (perception > memory > linguistic representation) production line, there is no evidence that any dorsal stream involvement is preserved at the stage of linguistic representations.

Perception of motion and mental representation of motion properties are at present probably the most problematic area for the central claim of the target article, and clearly more research, and perhaps some revision of the central claim, is necessary. But it is becoming clear that "motion" should not be treated as a single category. I mention below a few recent studies that suggest that at least some processing of motion takes place in the ventral stream. Beintema and Lappe (2002) report that "some patients with lesions to motion processing areas in the dorsal stream are severely impaired in image motion perception but can easily perceive biological motion" (p. 5661). Zhou et al. (2003) report that "Long-range AM [apparent motion] activated the anterior-temporal lobe in the visual ventral pathway, and the response varied according to form stability. The results suggest that long-range AM is associated with neural systems for form perception" (p. 417). Vaina et al. (2001) report "whereas face (and form) stimuli activate primarily the ventral system and motion stimuli primarily the dorsal system, recognition of biological motion stimuli may activate both systems as well as their confluence in STS."

R8. Binding, afference, and efference

Werning asks what, in my proposal, is the mechanism of binding an object concept to a property concept. (It would be closer to the concerns of the target article to ask about the binding of an object percept to a property percept, but that is a minor, perhaps terminological point.) The term *binding* is used in several contexts. The target article mentioned the "binding problem" at the end of section 2.2. This is the problem of how the brain represents the fact that several different properties belong to the same object. Werning mentions the "co-oscillation" solution, whereby neurons in anatomically connected regions registering different properties oscillate in synchrony if the properties belong to the same object. Given the insistence in the target article that objects are located by the dorsal stream and assigned properties by the ventral stream, a solution by co-oscillation in neighboring regions is not available to me, as Werning points out. **Bickerton** eloquently expresses the problem as follows:

> there must surely be some place in the brain for predicate and argument to come together. But on Hurford's account, there is nowhere for this to happen. One half of the predicate-argument equivalent occurs in the parietal cortex, the other half in the infero-temporal cortex. There would have to be efferent fibers from parietal to infero-temporal, or vice versa (or from both of these to some third place) if the two halves were to be integrated into either a thought or a sentence.

To this, Werning also says, "Hurford gives no answer." But I do, and it is in fact exactly what Bickerton claims as his own "more plausible (and more parsimonious) scenario," namely, that information from the dorsal stream alerts the organism to the fact that something of potential interest or importance is out there. Thereafter, it plays no direct role in cognition or language. The ventral stream carries richer information to (more or less) where concepts are stored. A match is made, or not, as the case may be. Efferent signals from parietal cortex direct gaze to the object, which allows information from that object to be transmitted via the afferent ventral stream. Bickerton's "some third place" is in a sense the perceived object itself. Didn't the target article put it plainly enough?

R9. Attention

Brinck focuses on the nature of attention. He first disagrees with the idea that objects of attention are "arbitrary." In fact, this term was only applied once to objects of attention, in the target article's abstract, and not used, implicitly or explicitly, in the body of the article. Nothing hinges on the word "arbitrary," and it should be withdrawn.

Brinck makes a valuable distinction, which I largely neglected, between stimulus-driven attention and goal-driven attention. As I understand Brinck's terminology, the process he calls "indexing" only applies in stimulus-driven attention. "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." I agree. Section 2.2 in the target article discussed "natural attention-drawing properties," as opposed to other kinds of properties. Brinck challenges this idea: "I do not see the need to introduce 'natural attention-drawing properties' to account for attention attraction." This seems inconsistent with the quotation above about indexing being caused by some property of the object. In his penultimate paragraph, Brinck writes that attention is attracted by sudden and unexpected changes in the subject's immediate environment. If such a sudden and unexpected change is to the *whole* environment, like the sudden darkness due to a total eclipse, or a bright light suddenly illuminating the whole of a previously dark room, then there is no single object to which attention is drawn. But if the change

is more locally constrained, almost certainly it will be a change in a property of some object, as seen from the subject's position. For example, a leaf may flutter or a door may open (I am happy with modes of movement being properties), or as the subject turns her head, redness appears, interpretable as some red object changing its position relative to the subject. Red is generally a more attention-drawing color than brown (which helps to account for the well-known hierarchy of Basic Color Terms in languages). The target article cited evidence that young children pay more attention to shape than to other properties of objects. It was largely stimulus-driven attention that was assumed in the target article, and I think the difference between red and brown makes the point. Some properties of objects grab attention faster and more effectively than others, and some properties of objects (such as their weight) hardly grab attention at all.

Turning now to goal-driven attention, it is only here that, as I understand **Brinck**'s terminology, one can speak of "targets of attention." "Goal-driven attention works top-down, in anticipation of some well-defined item. The subject is searching for a particular object." The target of attention is, then, the defining property of the sought-for object(s). So indexing is bottom-up, stimulus-driven, whereas having a target of attention happens in top-down, goal-driven search. Given this, Brinck is correct in saying that indexed objects can never be targets of attention. It follows from these definitions. To say otherwise would be like saying, contradictorily, "I'm looking for the thing that just immediately caught my attention."

The target article should have made the distinction between stimulus-driven and goal-driven attention. It was essentially about stimulus-driven attention. With that limitation, the arguments in the target article are not undermined by **Brinck**'s commentary. I suggest, furthermore, that stimulus-driven attention is the evolutionarily more primitive form of attention, thus rooting the neural basis of predicate-argument structure firmly in what **MacNeilage & Davis,** after Darwin, call "lowly origins."

R10. Action

Both **Indurkhya** and **MacNeilage & Davis** concentrate on action, rather than perception. MacNeilage & Davis emphasize that their account of the evolution of syllable structure, like mine of propositional structure, posits "lowly origins," that is, very ancient phylogenetic roots. They also emphasize the complementarity between their theory and mine, and Indurkhya's paper essentially presents a different choice of emphasis, rather than a refutation. Language, being a bridge between meanings and sounds, needs both semanticists and phoneticians. Unfortunately, semantics and phonetics are radically different disciplines, with entirely non-overlapping traditions of discourse. When a semanticist turns to thinking about the evolution of language, it is perhaps inevitable that he thinks about such matters as predicate-argument structure, and not syllable structure. Likewise, predicate-argument structure is far from the concerns of phoneticians.

I have much sympathy with the position of these writers that the evolutionary roots of language are to be found in action. "In the beginning was the deed, not the word," as Goethe's Faust insisted. The target article was mainly concerned with demonstrating a present-day correlation between semantic structure and neural organization. That this neural organization is shared by higher mammals does indicate "lowly origins," but I did not dwell on the evolutionary history of this organization (though it would be fascinating). At one point, I told a brief merely figurative story, repeated by **Indurkhya**, of the growth of predicate-argument structure from earlier forms of behavior which were holistic, and did not exhibit anything resembling the dichotomy between predicate and argument. I was once a phonetician, but it is too late for me to catch up with the likes of **MacNeilage & Davis** and theorize about the origins of speech. And if Indurkhya thinks that my story sped past the interesting bits too fast, he should write his own story.

Indurkhya raises the matter of holistic one-word utterances, as made by children and our ancestors at some stage. Only some such utterances support Indurkhya's view of an action-based system in which no division like that between subject and predicate can be made. If a speaker routinely grunts (like a tennis player) when performing a certain action, then certainly we may see the grunt as in some sense intrinsic to the action. But when a child says "Daddy!" as opposed to "Mummy," although the utterance is a single word, there are nevertheless distinguishable acts of referring to a particular person and assigning it a certain mental category. The target article noted briefly, near the end, that holistic utterances could nevertheless express predicate-argument meanings.

R11. Representations

I suspect that this topic is one on which the deepest divisions between researchers are to be found, reflecting fundamental metaphysical positions. In this section I sketch my own reductionist metaphysical position, and claim that it has the merit of parsimony.

Arbib makes what could seem to be an odd point about the distinction between neural processes and descriptions of those processes. Obviously, for any X, "description of X" is not the same as X. The word *electron* is not an electron. I agree with Arbib that the formula *PREDICATE(x)* is not itself a neural process. Who could think otherwise? Perhaps the issue is whether some neural process or configuration described by a scientist's predicate is itself a *representation* available to the animal concerned. I use *representation* in the sense that if an animal can reliably distinguish a certain class of stimuli from others, the neural configurations that enable it to do so *constitute* a representation of that class of stimuli, for which we humans may or may not happen to have a word, such as *red* or *leopard.* In this sense, the representation is available to the animal. I do not make the distinction between representations and "their supporting neuronal states and processes" made by **Piattelli-Palmarini & Harley**. For them, "representations are descriptions accessed internally by the subject."

It is an empirical matter what uses the animal can put its representations to. A frog can use its prey-representation for catching prey, but it cannot attach a symbolic label like *prey* or *insect* to its prey concept, for communicating about prey. Humans can describe their representations in a public code; most animals cannot. When a frog jumps at a particular stimulus, it would seem to be internally accessing (or