BEHAVIORAL AND BRAIN SCIENCES (2003) **26,** 261–316
Printed in the United States of America

The neural basis of predicate-argument structure

James R. Hurford
Language Evolution and Computation Research Unit, School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh EH8 9LL, Scotland, United Kingdom.
jim@ling.ed.ac.uk

Abstract: Neural correlates exist for a basic component of logical formulae, *PREDICATE(x).* Vision and audition research in primates and humans shows two independent neural pathways; one locates objects in body-centered space, the other attributes properties, such as colour, to objects. In vision these are the dorsal and ventral pathways. In audition, similarly separable "where" and "what" pathways exist. *PREDICATE(x)* is a schematic representation of the brain's integration of the two processes of delivery by the senses of the location of an arbitrary referent object, mapped in parietal cortex, and analysis of the properties of the referent by perceptual subsystems.

The brain computes actions using a few "deictic" variables pointing to objects. Parallels exist between such nonlinguistic variables and linguistic deictic devices. Indexicality and reference have linguistic and nonlinguistic (e.g., visual) versions, sharing the concept of attention. The individual variables of logical formulae are interpreted as corresponding to these mental variables. In computing action, the deictic variables are linked with "semantic" information about the objects, corresponding to logical predicates.

Mental scene descriptions are necessary for practical tasks of primates, and preexist language phylogenetically. The type of scene descriptions used by nonhuman primates would be reused for more complex cognitive, ultimately linguistic, purposes. The provision by the brain's sensory/perceptual systems of about four variables for temporary assignment to objects, and the separate processes of perceptual categorization of the objects so identified, constitute a pre-adaptive platform on which an early system for the linguistic description of scenes developed.

Keywords: argument; attention; deictic; dorsal; logic; neural; object; predicate; reference; ventral

1. Introduction

This article argues for the following thesis: Neural evidence exists for predicate-argument structure as the core of phylogenetically and ontogenetically primitive (prelinguistic) mental representations. The structures of modern natural languages can be mapped onto these primitive representations.

The idea that language is built onto preexisting representations is common enough, being found in various forms in works such as Bickerton (1998), Kirby (1999; 2000), Hurford (2000b), and Bennett (1976). Conjunctions of elementary propositions of the form *PREDICATE(x)* have been used by Batali as representations of conceptual structure preexisting language in his impressive computer simulations of the emergence of syntactic structure in a population of interacting agents (Batali 2002). Justifying such preexisting representations in terms of neural structure and processes is relatively new.

This paper starts from a very simple component of the Fregean logical scheme, *PREDICATE(x),* and proposes a neural interpretation for it. This is, to my knowledge, the first proposal of a "wormhole" between the hitherto mutually isolated universes of formal logic and empirical neuroscience. The fact that it is possible to show a correlation between neural processes and logicians' conclusions about logical form is a step in the unification of science. The discoveries in neuroscience confirm that the logicians have been on the right track, that the two disciplines have something to say to each other despite their radically different methods, and that further unification may be sought. As the brain has a complexity far in excess of any representation scheme dreamt up by a logician, it is to be expected that the basic *PREDICATE(x)* formalism is to some extent an idealization of what actually happens in the brain. But, conceding that the neural facts are messier than could be captured with absolute fidelity by any formula as simple as *PREDICATE(x),* I hope to show that the central ideas embodied in the logical formula map satisfyingly neatly onto certain specific neural processes.

The claim that some feature of language structure maps onto a feature of primitive mental representations needs (1) a plausible bridge between such representation and the structure of language, and (2) a characterization of

James R. Hurford is Professor of General Linguistics at the University of Edinburgh. He has a broad interest in reconciling various traditions in linguistics which have tended to conflict. In particular, he has worked on articulating a framework in which formal representation of grammars in individual minds interacts with properties of language as used in communities. The framework emphasizes the interaction of evolution, learning and communication. He is perhaps best known for his computer simulations of various aspects of the evolution of language.

"primitive mental representation" independent of language itself, to avoid circularity. The means of satisfying the first, the "bridge to language" condition, will be discussed in the next subsection. Fulfilling the second condition, the bridge to brain structure and processing, by establishing the language-independent validity of *PREDICATE(x)* as representing fundamental mental processes in both humans and nonhuman primates, will occupy the meat of this article (sects. 2 and 3). The article is original only in bringing together the fruits of others' labours. Neuroscientists andpsychologists will be familiar with much of the empirical research cited here, but I hope they will be interested in my claims for its wider significance. Linguists, philosophers, and logicians might be excited to discover a new light cast on their subjects by recent neurological research.

1.1. The bridge from logic to language

The relationship between language and thought is, of course, a vast topic, and there is only space here to sketch my premises about this relationship.

Descriptions of the structure of languages are couched in symbolic terms. Although it is certain that a human's knowledge of his/her language is implemented in neurons, and at an even more basic level of analysis, in atoms, symbolic representations are clearly well suited for the study of language structure. Neuroscientists don't need logical formulae to represent the structures and processes that they find. Ordinary language, supplemented by diagrams, mathematical formulae, and neologized technical nouns, verbs, and adjectives, is adequate for the expression of neuroscientists' amazingly impressive discoveries. Where exotic technical notations are invented, it is for compactness and convenience, and their empirical content can always be translated into more cumbersome ordinary language (with the technical nouns, adjectives, etc.).

Logical notations, on the other hand, were developed by scholars theorizing in the neurological dark about the structure of language and thought. Languages are systems for the expression of thought. The sounds and written characters, and even the syntax and phonology of languages can also be described in concrete ordinary language, augmented with diagrams and technical vocabulary. Here too, invented exotic notations are for compactness and convenience; which syntax lecturer has not paraphrased $S \rightarrow NP\ VP$ into ordinary English for the benefit of a first-year class? But the other end of the language problem, the domain of thoughts or meanings, has remained elusive to non-tautological ordinary language description. Of course, it is possible to use ordinary language to express thoughts – we do it all the time. But to say that "Snow is white" describes the thought expressed by "Snow is white" is either simply wrong (because description of a thought process and expression of a thought are not equivalent) or at best uninformative. To arrive at an informative characterization of the relation between thought and language (assuming the relation to be other than identity), you need some characterization of thought which does not merely mirror language. So logicians have developed special notations for describing thought (not that they have always admitted or been aware that that is what they were doing). But, up to the present, the only route that one could trace from the logical notations to any empirically given facts was *back through* the ordinary language expressions which motivated them in the first place. A neuroscientist can show you (using suitable instruments which you implicitly trust) the synapses, spikes, and neural pathways that he investigates. But the logician cannot illuminatingly bring to your attention the logical form of a particular natural sentence, without using the sentence itself, or a paraphrase of it, as an instrument in his demonstration. The mental adjustment that a beginning student of logic is forced to make, in training herself to have the "logician's mindset," is absolutely different in kind from the mental adjustment that a beginning student of a typical empirical science has to make. One might, prematurely, conclude that logic and the empirical sciences occupy different universes, and that no wormhole connects them.

Despite its apparently unempirical character, logical formalism is not mere arbitrary stipulation, as some physical scientists may be tempted to believe. One logical notation can be more explanatorily powerful than another, as Frege's advances show. Frege's introduction of quantifiers binding individual variables which could be used in argument places was a great leap forward from the straitjacket of subject-predicate structure originally proposed by Aristotle and not revised for over two millennia. Frege's new notation (but not its strictly graphological form, which was awfully cumbersome) allowed one to explain thoughts and inferences involving a far greater range of natural sentences. Logical representations, systematically mapped to the corresponding sentences of natural languages, clarify enormously the system underlying much human reasoning, which, without the translation to logical notation, would appear utterly chaotic and baffling.

It is necessary to note a common divergence of usage, between philosophers and linguists, in the term "subject." For some philosophers (e.g., Strawson 1959; 1974), a predicate in a simple proposition, as expressed by *John loves Mary,* for example, can have more than one "subject;" in the example given, the predicate corresponds to *loves* and its "subjects" to *John* and *Mary.* In this usage, the term "subject" is equivalent to "argument." Linguists, on the other hand, distinguish between grammatical subjects and grammatical objects, and further between direct and indirect objects. Thus, in *Russia sold Alaska to America,* the last two nouns are not subjects, but direct and indirect object respectively. The traditional grammatical division of a sentence into Subject + Predicate is especially problematic where the "Predicate" contains several NPs, semantically interpreted as arguments of the predicate expressed by the verb. Which argument of a predicate, if any, is privileged to be expressed as the grammatical subject of a sentence (thus in English typically occurring before the verb, and determining number and person agreement in the verb) is not relevant to the truth-conditional analysis of the sentence. Thus, a variety of sentences such as *Alaska was sold to America by Russia* and *It was America that was sold Alaska by Russia* all describe the same state of affairs as the earlier example. The difference between the sentences is a matter of rhetoric, or appropriate presentation of information in various contextual circumstances, involving what may have been salient in the mind of the hearer or reader before encountering the sentence, or how the speaker or writer wishes to direct the subsequent discourse.

Logical predicates are expressed in natural language by words of various parts of speech, including verbs, adjectives, and common nouns. In particular, there is no special connection between grammatical verbs and logical predicates. The typical correspondences between the main English syntactic categories and basic logical terms are diagrammed below.

	Pronouns	– – – Arguments
Nouns —	Proper Names	
	Common Nouns	– – – 1, 2, . . . n–place Predicates
	Verbs	
	Adjectives	
	Prepositions	
	Adverbs	

Common nouns, used after a copula, as *man* in *He is a man* plainly correspond to predicates. In other positions, although they are embedded in grammatical noun phrases, as in *A man arrived,* they nonetheless correspond to predicates.

The development of formal logical languages, of which first order predicate logic is the foremost example and hardiest survivor, heralds a realization of the essential distance between ordinary language and purely truth-conditional representations of "objective" situations in the world. Indeed, early generations of modern logicians, including Frege, Russell, and Tarski, believed the gap between ordinary language and logical, purely truth-conditional representations to be unbridgable. Times have changed, and since Montague there have been substantial efforts to describe a systematic mapping between truth conditions and ordinary language. Ordinary language serves several purposes in addition to representation of states of affairs. My argument in this article concerns mental representations of situations in the world, as these representations existed *before* language, and even before communication. Thus, matters involving how information is presented in externalized utterances are not our concern here. The exclusive concern here with pre-communication mental representations absolves us from responsibility to account for further cognitive properties assumed by more or less elaborate signals in communication systems, such as natural languages. For this reason also, the claims to be made here about the neural correlates of *PREDICATE(x)* do not relate at all directly to matters of *linguistic* processing (e.g., sentence parsing), as opposed to the prelinguistic representation of events and situations.

Bertrand Russell was, of course, very far from conceiving of the logical enterprise as relating to how non-linguistic creatures represent the world. But it might be helpful to note that Russell's kind of flat logical representations, as in $\exists x$ *[KoF(x) & wise(x)]* for *The king of France is wise,*[1] are essentially like those assumed by Batali (2002) and focussed on in this article. Russell's famous controversy with Strawson (Russell 1905; 1957; Strawson 1950) centered on the effect of embedding an expression for a predicate in a noun phrase determined by the definite article. Questions of definiteness only arise in communicative situations, with which Strawson was more concerned. A particular object in the world is inherently neither definite nor indefinite; only when we talk about an object do our referring noun phrases begin to have markers of definiteness, essentially conveying "You are already aware of this thing."

The thesis proposed here is that there were, and still are, pre-communication mental representations which embody the fundamental distinction between predicates and arguments, and in which the foundational primitive relationship is that captured in logic by formulae of the kind *PREDICATE(x).* The novel contribution here is that the centrality of predicate-argument structure has a neural basis, adapted to a sentient organism's traffic with the world, rather than having to be postulated as "logically true" or even Platonically given. Neuroscience can, I claim, offer some informative answers to the question of where elements of logical form came from.

The strategy here is to assume that a basic element of first order predicate logic notation, *PREDICATE(x),* suitably embedded, can be systematically related to natural language structures, in the ways pursued by recent generations of formal semanticists of natural language, for example, Montague (1970; 1973), Parsons (1990), Kamp and Reyle (1993). The hypothesis here is *not* that *all* linguistic structure derives from prelinguistic mental representations. I argue elsewhere (Hurford 2002) that in fact very little of the rich structure of modern languages directly mirrors any mental structure in pre-existing language.

In generative linguistics, such terms as "deep structure" and "surface structure," "logical form" and "phonetic form" have specialized theory-internal meanings, but the basic insight inherent in such terminology is that linguistic structure is a *mapping between* two distinct levels of representation. In fact, most of the complexity in language structure belongs to this mapping, rather than to the forms of the anchoring representations themselves. In particular, the syntax of logical form is very simple. All of the complexities of phonological structure belong to the mapping between meaning and form, rather than to either meaning or form *per se.* A very great proportion of morphosyntactic structure clearly also belongs to this mapping – components such as word-ordering, agreement phenomena, anaphoric marking, most syntactic category distinctions (e.g., noun, verb, auxiliary, determiner), which have no counterparts in logic, and focussing and topicalization devices. In this respect, the view taken here differs significantly from Bickerton's (in Calvin & Bickerton 2000) that modern grammar in all its glory can be derived, with only a few auxiliary assumptions, from the kind of mental representations suitable for cheater detection that our prelinguistic ancestors would have been equipped with; see Hurford (2002) for a fuller argument.

Therefore, to argue, as I will in this article, that a basic component of the representation of meaning preexists language and can be found in apes, monkeys, and possibly other mammals, leaves most of the structure of language (the complex mappings of meanings to phonetic signals) still unexplained in evolutionary terms. To argue that apes have representations of the form *PREDICATE(x),* does not make them out to be language-capable humans. Possession of the *PREDICATE(x)* form of representation is evidently not sufficient to propel a species into full-blown syntactic language. There is much more to human language than predicate-argument structure, but predicate-argument structure is the semantic foundation on which all the rest is built.

The view developed here is similar in its overall direction to that taken by Bickerton (1990). Bickerton argues for a "primary representation system (PRS)" existing in variously developed forms in all higher animals. "In all probability, language served in the first instance merely to label protoconcepts derived from prelinguistic experience" (p. 91). This is entirely consistent with the view proposed here, assuming that what I call "prelinguistic mental predicates" are Bickerton's "protoconcepts." Bickerton also believes, as I do, that the representation systems of prelinguistic creatures have predicate-argument structure. Bickerton further suggests that, even before the emergence of language, it is possible to distinguish subclasses of mental predicates along lines that will eventually give rise to linguistic distinctions such as Noun/Verb. He argues that "[concepts corresponding to] verbs are much more abstract than [those corresponding to] nouns" (p. 98). I also believe that a certain basic functional classification of predicates can be argued to give rise to the universal linguistic categories of Noun and Verb. But that subdivision of the class of predicates is not my concern here. Here the focus is on the more fundamental issue of the distinction between predicates and their arguments. So this paper is *not* about the emergence of Noun/Verb structure (which is a story that must wait for another day). (Batali's [2002] impressive computer simulations of the emergence of some aspects of natural language syntax start from conjunctions of elementary formulae in *PREDICATE(x)* form, but it is notable that they do not arrive at anything corresponding to a Noun/Verb distinction.)

On top of predicate-argument structure, a number of other factors need to come together for language to evolve. Only the sketchiest mention will be given of such factors here, but they include (a) the transition from private mental representations to public signals; (b) the transition from involuntary to voluntary control; (c) the transition from epigenetically determined to learned and culturally transmitted systems; (d) the convergence on a common code by a community; (e) the evolution of control of complex hierarchically organized signalling behaviour (syntax); (f) the development of deictic here-and-now talk into definite reference and proper naming capable of evoking events and things distant in time and space. It is surely a move forward in explaining the evolution of language to be able to dissect out the separate steps that must be involved, even if these turn out to be more dauntingly numerous than was previously thought. (In parallel fashion, the discovery of the structure of DNA immediately posed problems of previously unimagined complexity to the next generation of biologists.)

1.2. Prelinguistic predicates

In the view adopted here, a predicate corresponds, to a first approximation, to a judgement that a creature can make about an object. Some predicates are relatively simple. For a simple predicate, the senses provide the brain with input allowing a decision with relatively little computation. On a scale of complexity, basic colour predicates are near the simple end, while predicates paraphrasable as *sycamore* or *weasel* are much more complex. Mentally computing the applicability of complex predicates often involves simpler predicates, hence relatively more computation.

Some ordinary languages predicates, such as *big*, depend for their interpretation on the prior application of other predicates. Generically speaking, a big flea is not big; this is no contradiction, once it is admitted that the sentence implicitly establishes two separate contexts for the application of the adjective *big.* There is "big, generically speaking," that is, in the context of consideration of all kinds of objects and of no one kind of object in particular; and there is "big for a flea." This is semantic modulation. Such modulation is not a solely linguistic phenomenon. Many of our higher-level perceptual judgements are modulated in a similar way. An object or substance characterized by its whitish colour (like chalk) reflects bright light in direct sunlight, but a light of lower intensity in the shade at dusk. Nevertheless, the brain, in both circumstances, is able to categorize this colour as whitish, even though the lower intensity of light is reflected by a greyish object or substance (like slate) in direct sunlight. In recognizing a substance as whitish or greyish, the brain adjusts to the ambient lighting environment. Viewing chalk in poor light, the visual system returns the judgement "Whitish, for poor light"; in response to light of the same intensity, as when viewing slate in direct sunlight, the visual system returns the judgement "Greyish, for broad daylight." A similar example can be given from speech perception. In a language such as Yoruba, with three level lexical tones, high, mid, and low, a single word spoken by an unknown speaker cannot reliably be recognized as on a high tone spoken by a man or a low or mid tone spoken by a woman or child. But as soon as a few words are spoken, the hearer recognizes the appropriate tones in the context of the overall pitch range of the speaker's voice. Thus, the ranges of external stimuli which trigger a mental predicate may vary, systematically, as a function of other stimuli present.

This article will be mainly concerned with 1-place predicates, arguing that they correspond to perceived properties. There is no space here to present a fully elaborated extension of the theory to predicates of degree greater than one, but a few suggestive remarks may convince a reader that in principle the theory may be extendable to *n*-place predicates ($n > 1$).

Prototypical events or situations involving 2-place predicates are described by *John kicked Fido* (an event) or *The cat is on the mat* (a situation). Here I will take it as given that observers perceive events or situations as unified wholes; there is some psychological reality to the concept of an atomic event or situation. In a 2-place predication (barring predicates used reflexively), the two participant entities involved in the event or situation also have properties. In formal logic, it is possible to write a formula such as $\exists x\ \exists y\ [kick(x, y)]$, paraphrasable as *Something kicks something.* But I claim that it is never possible for an observer to perceive an event of this sort without also being able to make some different 1-place judgements about the participants. Perhaps the most plausible potential counterexample to this claim would be reported as *I feel something.* Now this could be intended to express a 1-place state, as in *I am hungry;* but if it is genuinely intended as a report of an experience involving an entity other than the experiencer, I claim that there will always be some (1-place) property of this entity present to the mind of the reporter. That is, the "something" which is felt will always be felt as having some property, such as sharpness, coldness or furriness. Expressed in terms of a psychologically realistic logical language enhanced by meaning postulates, this amounts to the claim that every 2-place predicate occurs in the *implicans* of some meaning postulate whose *implicatum* includes 1-place predicates applicable to its arguments. The selectional restrictions expressed in some generative grammars provide good examples; the subject of *drink* must be animate, the object of *drink* must be a liquid.

In the case of asymmetric predicates, the asymmetry can always be expressed in terms of one participant in the event or situation having some property which the other lacks. And, I suggest, this treatment is psychologically plausible. In cases of asymmetric actions, as described by such verbs as *hit* and *eat,* the actor has the metaproperty of being the actor, cashed out in more basic properties such as movement, animacy, and appearance of volition. Likewise, the other, passive, participant is typically characterized by properties such as lack of movement, change of state, inanimacy, and so forth (see Cruse 1973 and Dowty 1991 for relevant discussion). Cases of asymmetric situations, such as are involved in spatial relations as described by prepositions such as *on, in,* and *under,* are perhaps less obviously treatable in this way. Here, I suggest that properties involving some kind of perceptual salience in the given situation are involved. In English, while both sentences are grammatical, *The pen is on the table* is commonplace, but *The table is under the pen* is studiously odd. I would suggest that an object described by the grammatical subject of *on* has a property of being taken in as a whole object comfortably by the eye, whereas the other object involved lacks this property and is perceived (on the occasion concerned) rather as a *surface* than as a whole object.

In the case of symmetric predicates, as described by *fight each other* or *as tall as,* the arguments are not necessarily distinguished by any properties perceived by an observer.

I assume a version of event theory (Davidson 1980; Parsons 1990), in which the basic ontological elements are whole events or situations, annotated as *e,* and the participants of these events, typically no more than about three, annotated as *x, y,* and *z.* For example, the event described by *A man bites a dog* could be represented as $\exists e, x, y, bite(e), man(x), dog(y), agent(x), patient(y)$. In clumsy English, this corresponds to "There is a biting event involving a man and a dog, in which the man is the active volitional participant, and the dog is the passive participant." The less newsworthy event would be represented as $\exists e, x, y, bite(e), man(x), dog(y), agent(y), patient(x)$. The situation described by *The pen is on the table* could be represented as $\exists e, x, y, on(e), pen(x), table(y), small\_object(x), surface(y)$.

In this enterprise it is important to realize the great ambiguity of many ordinary language words. The relations expressed by English *on* in *An elephant sat on a tack* and in *A book lay on a table* are perceptually quite different (though they also have something in common). Thus, there are at least several mental predicates corresponding to ordinary language words. When, in the histories of natural languages, words change their meanings, the overt linguistic forms become associated with *different* mental predicates. The predicates which I am concerned with here are prelinguistic mental predicates, and are not to be simply identified with words.

Summarizing these notes, it is suggested that it may be possible to sustain the claim that n-place predicates ($n > 1$) are, at least in perceptual terms, constructible from 1-place predicates. The core of my argument in this article concerns formulae of the form *PREDICATE(x),* that is, 1-place predications. My core argument in this article does not stand or fall depending on the correctness of these suggestions about $n > 1$-place predicates. If the suggestions about $n > 1$-place predicates are wrong, then the core claim is limited to 1-place predications, and some further argument will need to be made concerning the neural basis of $n > 1$-place predications. A unified theory relating all logical predicates to the brain is methodologically preferable, so there is some incentive to pursue the topic of $n > 1$-place predicates.

1.3. Individual variables as prelinguistic arguments

Here are two formulae of first order predicate logic (FOPL), with their English translations.

$CAME(john)$ (Translation: "John came")
$\exists x[TALL(x) \& MAN(x) \& CAME(x)]$ (Translation: "A tall man came")

The canonical fillers of the argument slots in predicate logic formulae are constants denoting individuals, corresponding roughly to natural language proper names. In the more traditional schemes of semantics, no distinction between extension and intension is made for proper names. On many accounts, proper names have only extensions (namely the actual individuals they name), and do not have intensions (or "senses"). "What is probably the most widely accepted philosophical view nowadays is that they [proper names] may have reference, but not sense" (Lyons 1977, p. 219). "Dictionaries do not tell us what [proper] names mean – for the simple reason that they do not mean anything" (Ryle 1957). In this sense, the traditional view has been that proper names are semantically simpler than predicates. More recent theorizing has questioned that view.

In a formula such as *CAME(john),* the individual constant argument term is interpreted as denoting a particular individual, the very same person on all occasions of use of the formula. FOPL stipulates by fiat this absolutely fixed relationship between an individual constant and a particular individual entity. Note that the denotation of the term is a thing in the world, outside the mind of any user of the logical language. It is argued at length by Hurford (2001) that the mental representations of protohumans could not have included terms with this property. Protothought had no equivalent of proper names. Control of a proper name in the logical sense requires Godlike omniscience. Creatures only have their sense organs to rely on when attempting to identify, and to re-identify, particular objects in the world. Where several distinct objects, identical to the senses, exist, a creature cannot reliably tell which is which, and therefore cannot guarantee control of the fixed relation between an object and its proper name that FOPL stipulates. It's no use applying the same name to each of them, because that violates the requirement that logical languages be unambiguous. More detailed arguments along these lines are given in Hurford (1999; 2001), but it is worth repeating here the counterargument to the most common objection to this idea. It is commonly asserted that animals can recognize other animals in their groups.

The following quotation demonstrates the *prima facie* attraction of the impression that animals distinguish such individuals, but simultaneously gives the game away.

> The speed with which recognition of individual parents can be acquired is illustrated by the "His Master's Voice" experiments performed by Stevenson et al. (1970) on young terns: these responded immediately to tape-recordings of their own parents (by cheeping a greeting, and walking towards the loudspeaker) but ignored other tern calls, even those recorded from other adult members of their own colony. (Walker 1983, p. 215)

Obviously, the tern chicks in the experiment were *not* recognizing their individual parents – they were being fooled into treating a loudspeaker as a parent tern. For the tern chick, anything that behaved sufficiently like its parent was "recognized" as its parent, even if it wasn't. The tern chicks were responding to very finely-grained *properties* of the auditory signal, and apparently neglecting even the most obvious of visual properties discernible in the situation. In tern life, there usually aren't human experimenters playing tricks with loudspeakers, and so terns have evolved to discriminate between auditory cues just to the extent that they can identify their own parents with a high degree of reliability. Even terns presumably sometimes get it wrong.

> Animals respond in mechanical robot-like fashion to key stimuli. They can usually be "tricked" into responding to crude dummies that resemble the true, natural stimulus situation only partially, or in superficial respects. (Krebs & Dawkins 1984, p. 384; quoted in Hurford 2001)

The logical notion of an individual constant permits no degree of tolerance over the assignment of these logical constants to individuals; this is why they are called "constants." It is an *a priori* fiat of the design of the logical language that individual constants pick out particular individuals with absolute consistency. In this sense, the logical language is practically unrealistic, requiring, as previously mentioned, Godlike omniscience on the part of its users, the kind of omniscience reflected in the biblical line "But even the very hairs of your head are all numbered" (Matthew, Ch.10).

Interestingly, several modern developments in theorizing about predicates and their arguments complicate the traditional picture of proper names, the canonical argument terms. The dominant analysis in the modern formal semantics of natural languages (e.g., Montague 1970; 1973) does not treat proper names in languages (e.g., *John*) like the individual constants of FOPL. For reasons having to do with the overall generality of the rules governing the compositional interpretation of all sentences, modern logical treatments make the extensions of natural language proper names actually *more complex than,* for example, the extensions of common nouns, which are 1-place predicates. In such accounts, the extension of a proper name is not simply a particular entity, but the set of classes containing that entity, while the extension of a 1-place predicate is a class. Concretely, the extension of *cat* is the class of cats, while the extension of *John* is the set of all classes containing John.

Further, it is obvious that in natural languages, there are many kinds of expressions other than proper names which can fill the NP slots in clauses.

Semantically, then, PNs are an incredibly special case of NP; almost nothing that a randomly selected full NP can denote is also a possible proper noun denotation. This is surprising, as philosophers and linguists have often treated PNs as representative of the entire class of NPs. Somewhat more exactly, perhaps, they have treated the class of full NPs as representable . . . by what we may call *individual denoting* NPs (Keenan 1987, p. 464).

This fact evokes one of two responses in logical accounts. The old-fashioned way was to deny that there is any straightforward correspondence between natural language clauses with nonproper name subjects or objects and their translations in predicate logic (as Russell [1905] did). The modern way is to complicate the logical account of what grammatical subjects (and objects), including proper names, actually denote (as Montague did).

In sum, logical formulae of the type *CAME(john),* containing individual constants, cannot be plausibly claimed as corresponding to primitive mental representations preexisting human language. The required fixing of the designations of the individual constants ("baptism" in Kripke's [1980] terms) could not be practically relied upon. Modern semantic analysis suggests that natural language proper names are in fact more complex than longer noun phrases like *the man,* in the way they fit into the overall compositional systems of modern languages. And while proper names provide the shortest examples of (nonpronominal) noun phrases, and hence are convenient for brief expository examples, they are in fact somewhat peripheral in their semantic and syntactic properties.

Such considerations suggest that, far from being primitive, proper names are more likely to be relatively late developments in the evolution of language. In the historical evolution of individual languages, proper names are frequently, and perhaps always, derived from definite descriptions, as is still obvious from many, such as, *Baker, Wheeler, Newcastle.* It is very rare for languages to lack proper names, but such languages do exist. Machiguenga (or Matsigenka), an Arawakan language, is one, as several primary sources (Johnson 2003; Snell 1964) testify.

A most unusual feature of Matsigenka culture is the near absence of personal names (Snell 1964, pp. 17–25). Since personal names are widely regarded by anthropologists as a human universal (e.g., Murdock 1960, p. 132), this startling assertion is likely to be received with skepticism. When I first read Snell's discussion of the phenomenon, before I had gone into the field myself, I suspected that he had missed something (perhaps the existence of secret ceremonial names) despite his compelling presentation of evidence and his conclusion:

> I have said that the names of individual Machiguenga, when forthcoming, are either of Spanish origin and given to them by the white man, or nicknames. We have known Machiguenga Indians who reached adulthood and died without ever having received a name or any other designation outside of the kinship system. . . . Living in small isolated groups there is no imperative need for them to designate each other in any other way than by kinship terminology. Although there may be only a few tribes who do not employ names. I conclude that the Machiguenga is one of those few. (Snell 1964, p. 25)

Experience has taught me that Snell was right. Although the Matsigenka of Shimaa did learn the Spanish names given them, and used them in instances where it was necessary to refer to someone not of their family group, they rarely used them otherwise and frequently forgot or changed them. (Johnson 2003)

Joseph Henrich, another researcher on Machiguenga tells me "This is a well established fact among Machiguenga researchers" (personal communication).

> In this society there is very little cooperation, exchange or sharing beyond the family unit. This insularity is reflected in the fact that until recently they didn't even have personal names, referring to each other simply as "father," "patrilineal same-sex cousin," or whatever. (Douglas 2001, p. 41)

The social arrangements of our prelinguistic ancestors probably involved no cooperation, exchange, or sharing beyond the family unit, and the mental representations which they associated with individuals could well have been kinship predicates or other descriptive predicates.

In Australian languages, people are usually referred to by descriptive predicates.

> Each member of a tribe will also have a number of personal names, of different types. They may be generally known by a nickname, describing some incident in which they were involved or some personal habit or characteristic, e.g., "[she who] knocked the hut over," "[he who] sticks out his elbows when walking," "[she who] runs away when a boomerang is thrown," "[he who] has a damaged foot." But each individual will also have a sacred name, generally given soon after birth. (Dixon 1980, p. 27)

The extensive anthropological literature on names testifies to the very special status, in a wide range of cultures, of such sacred or "baptismal" proper names, both for people and places. It is common for proper names to be used with great reluctance, for fear of giving offense or somehow intruding on a person's mystical selfhood. A person's proper name is sometimes even a secret.

> The personal names by which a man is known are something more than names. Native statements suggest that names are thought to partake of the personality which they designate. The name seems to bear much the same relation to the personality as the shadow or image does to the sentient body. (Stanner 1937, quoted in Dixon 1980, p. 28)

It is hard to see how such mystical beliefs can have become established in the minds of creatures without language. More probably, it was only early forms of language itself that made possible such elaborate responses to proper names.

Hence, it is unlikely that any primitive mental representation contained any equivalent of a proper name, that is, an individual constant. We thus eliminate formulae of the type of *CAME(john)* as candidates for primitive mental representations.

This leaves us with quantified formulae, as in $\exists x$ *[MAN(x) & TALL(x)].* Surely we can discount the universal quantifier $\forall$ as a term in primitive mental representations. What remains is one quantifier, which we can take to be implicitly present and to bind the variable arguments of predicates. I propose that formulae of the type *PREDICATE(x)* are evolutionarily primitive mental representations, for which we can find evidence outside language.

2. Neural correlates of PREDICATE(x)

It is high time to mention the brain. In terms of neural structures and processes, what justification is there for positing representations of the form *PREDICATE(x)* inside human heads? I first set out some ground rules for correlating logical formulae, defined denotationally and syntactically, with events in the brain.

Representations of the form *PREDICATE(x)* are here interpreted psychologistically; specifically, they are taken to stand for the mental events involved when a human attends to an object in the world and classifies it perceptually as satisfying the predicate in question. In this psychologistic view, it seems reasonable to correlate *denotation* with *stimulus.* Denotations belong in the world outside the organism; stimuli come from the world outside a subject's head. A whole object, such as a bird, can be a stimulus. Likewise, the properties of an object, such as its colour or shape, can be stimuli. The two types of term in the *PREDICATE(x)* formula differ in their denotations. An individual variable does not have a constant denotation, but is assigned *different denotations on different occasions of use;* and the denotation assigned to such a variable is some *object* in the world, such as a particular bird, or a particular stone or a particular tree. A predicate denotes a *constant property* observable in the world, such as greenness, roundness, or the complex property of being a certain kind of bird. The question to be posed to neuroscience is whether we can find separate neural processes corresponding to (1) the shifting, ad hoc assignment of a "mental variable" to different stimulus objects in the world, not necessarily involving all, or even many, of the objects' properties, and (2) the categorization of objects, once they instantiate mental object variables, in terms of their properties, including more immediate perceptual properties, such as colour, texture, and motion, and more complex properties largely derived from combinations of these.

The syntactic structure of the *PREDICATE(x)* formula combines the two types of term into a unified whole capable of receiving a single interpretation which is a function of the denotations of the parts; this whole is typically taken to be an event or a state of affairs in the world. The bracketing in the *PREDICATE(x)* formula is not arbitrary: it represents an asymmetric relationship between the two types of information represented by the variable and the predicate terms. Specifically, the predicate term is understood in some sense to operate on, or apply to, the variable, whose value is provided beforehand. The bracketing in the *PREDICATE(x)* formula is the first, lowest-level, step in the construction of complex hierarchical semantic structures, as provided, for example, in more complex formulae of FOPL. The innermost brackets in a FOPL formula are always those separating a predicate from its arguments. If we can find separate neural correlates of individual variables and predicate constants, then the question to be put to neuroscience about the validity of the whole formula is whether the brain actually at any stage applies the predicate (property) system to the outputs of the object variable system, in a way that can be seen as the bottom level of complex, hierarchically organized brain activity.

2.1. Separate locating and identifying components in vision and hearing

The evidence cited here is mainly from vision. Human vision is the most complex of all sensory systems. About a quarter of human cerebral cortex is devoted to visual analysis and perception. There is more research on vision relevant to our theme, but some work on hearing has followed the recent example of vision research and arrived at similar conclusions.

2.1.1. Dorsal and ventral visual streams. Research on the neurology of vision over the past two decades has reached two important broad conclusions. One important message from the research is that vision is not a single unified system: Perceiving an object as having certain properties is a complex process involving clearly distinguishable pathways, and hence processes, in the brain (seminal works are Goodale & Milner 1992; Trevarthen 1968; Ungerleider & Mishkin 1982).

The second important message from this literature, as argued, for instance, by Milner and Goodale (1995), is that much of the visual processing in any organism is inextricably linked with motor systems. If we are to carve nature at her joints, the separation of vision from motor systems is in many instances untenable. For many cases, it is more reasonable to speak of a number of visuomotor systems. Thus, frogs have distinct visuomotor systems for orienting to and snapping at prey, and for avoiding obstacles when jumping (Ingle 1973; 1980; 1982). Distinct neural pathways from the frog's retina to different parts of its brain control these reflex actions.

Distinct visuomotor systems can similarly be identified in mammals, as Milner and Goodale (1995) report:

> In summary, the modular organization of visuomotor behaviour in representative species of at least one mammalian order, the rodents, appears to resemble that of much simpler vertebrates such as the frog and toad. In both groups of animals, visually elicited orienting movements, visually elicited escape, and visually guided locomotion around barriers are mediated by quite separate pathways from the retina right through to motor nuclei in the brainstem and spinal cord. This striking homology in neural architecture suggests that modularity in visuomotor control is an ancient (and presumably efficient) characteristic of vertebrate brains. (Milner & Goodale 1995, pp. 18–19)

Coming closer to our species, a clear consensus has emerged in primate (including human) vision research that one must speak of (at least) two separate neural pathways involved in the vision-mediated perception of an object. The literature is centred around discussion of two related distinctions: the distinction between magno and parvo channels from the retina to the primary visual cortex (V1) (Livingstone & Hubel 1988), and the distinction between dorsal and ventral pathways leading from V1 to further visual cortical areas (Mishkin et al. 1983; Ungerleider & Mishkin 1982). These channels and pathways function largely independently, although there is some crosstalk between them (Merigan et al. 1991; Van Essen et al. 1992), and in matters of detail there is, naturally, complication (e.g., Hendry & Yoshioka 1994; Johnsrude et al. 1999; Marois et al. 2000) and some disagreement (e.g., Franz et al. 2000; Merigan & Maunsell 1993; Zeki 1993). See Milner and Goodale (1995, pp. 33–39, 134–36) for discussion of the magno/parvo-dorsal/ventral relationship. (One has to be careful what one understands by "modular" when quoting Milner & Goodale [1995]. In real brains, modules are neural entities that modulate, compete, and cooperate, rather than being encapsulated processors for one "faculty" [Arbib 1987b].) It will suffice here to collapse under the label "dorsal stream" two separate pathways from the retina to posterior parietal cortex; one route passes via the lateral geniculate nucleus and V1, and the other bypasses V1 entirely, passing through the superior colliculus and pulvinar (see Milner & Goodale 1995, p. 68).

While it is not obvious that both divergences pertain to the same functional role, the proposals made here are not so detailed or subtle as to suggest any relevant discrimination between these two branches of the route from retina to parietal cortex. The dorsal stream has been characterized as the "where" stream, and the ventral stream as the "what" stream. The popular "where" label can be misleading, suggesting a single system for computing all kinds of spatial location; as we shall see, a distinction must be made between the computing of egocentric (viewer-centred) locational information and allocentric (other-centred) locational information. Bridgeman et al. (1979) use the preferable terms "cognitive" (for "what" information) and "motor-oriented" (for "where" information). Another suitable mnemonic might be the "looking" stream (dorsal) and the "seeing" stream (ventral). Looking is a visuomotor activity, involving a subset of the information from the retina controlling certain motor responses such as eye-movement, head and body orientation, and manual grasping or pointing. Seeing is a perceptual process, allowing the subject to deploy other information from the retina to ascribe certain properties, such as colour and motion, to the object to which the dorsal visuomotor looking system has already directed attention.

> Appreciation of an object's qualities and of its spatial location depends on the processing of different kinds of visual information in the inferior temporal and posterior parietal cortex, respectively. (Ungerleider & Mishkin 1982, p. 578)

Both cortical streams process information about the intrinsic properties of objects and their spatial locations, but the transformations they carry out reflect the different purposes for which the two streams have evolved. The transformations carried out in the ventral stream permit the formation of perceptual and cognitive representations which embody the enduring characteristics of objects and their significance; those carried out in the dorsal stream, which need to capture instead the instantaneous and egocentric features of objects, mediate the control of goal-directed actions (Milner & Goodale 1995, pp. 65–66).

Figure 1 shows the separation of dorsal and ventral pathways in schematic form.

Experimental and pathological data support the distinction between visuo-perceptual and visuomotor abilities.

Patients with cortical blindness, caused by a lesion to the visual cortex in the occipital lobe, sometimes exhibit "blindsight." Sometimes the lesion is unilateral, affecting just one hemifield, sometimes bilateral, affecting both; presentation of stimuli can be controlled experimentally, so that conclusions can be drawn equally for partially and fully blind patients. In fact, paradoxically, patients with the blindsight condition are never strictly "fully" blind, even if both hemifields are fully affected. Such patients verbally disclaim ability to see presented stimuli, and yet they are able to carry out precisely guided actions such as eye-movement, manual grasping and "posting" (into slots). (See Goodale et al. 1994; Marcel 1998; Milner & Goodale 1995; Sanders et al. 1974; Weiskrantz 1986; 1997. See also Ramachandran & Blakeslee [1998] for a popular account.)

These cited works on blindsight conclude that the spared unconscious abilities in blindsight patients are those identifying relatively low-level features of a "blindly seen" object, such as its size and distance from the observer, while access to relatively higher-level features such as colour and some aspects of motion is impaired.[2] Classic blindsight cases arise with humans, who can report verbally on their inability to see stimuli, but parallel phenomena can be tested and observed in nonhumans. Moore et al. (1998)

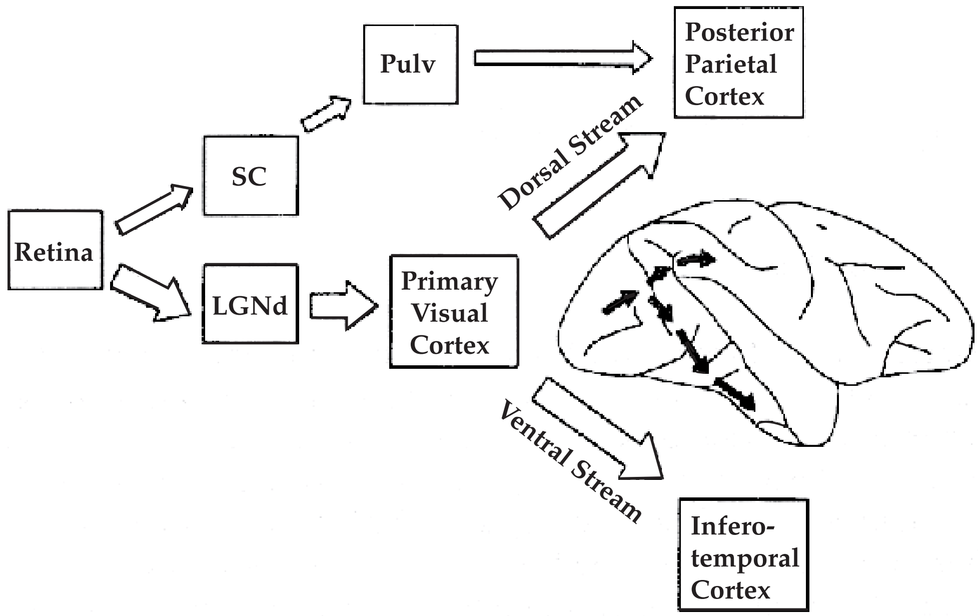

Figure 1. Schematic diagram showing major routes whereby retinal input reaches dorsal and ventral streams. The inset [brain drawing] shows the cortical projections on the right hemisphere of a macaque brain. LGNd, lateral geniculate nucleus, pars dorsalis; Pulv, pulvinar nucleus; SC, superior colliculus (From Milner & Goodale 1995).

summarize parallels between residual vision in monkeys and humans with damage to V1.

A converse to the blindsight condition has also been observed, indicating a double dissociation between visually-directed grasping and visual discrimination of objects. Goodale et al.'s patient R.V. could discriminate one object from another, but was unable to use visual information to grasp odd-shaped objects accurately (Goodale et al. 1994). Experiments with normal subjects also demonstrate a mismatch between verbally reported visual impressions of the comparative size of objects and visually-guided grasping actions. In these experiments, subjects were presented with a standard size-illusion-generating display, and asserted (incorrectly) that two objects differed in size; yet when asked to grasp the objects, they spontaneously placed their fingers exactly the same distance apart for both objects (Aglioti et al. 1995). Aglioti et al.'s conclusions have recently been called into question by Franz et al. (2000); see the discussion by Westwood et al. (2000) for a brief up-to-date survey of nine other studies on this topic.

Advances in brain-imaging technology have made it possible to confirm in nonpathological subjects the distinct localizations of processing for object recognition and object location (e.g., Aguirre & D'Esposito [1997] and other studies cited in this paragraph). Haxby et al. (1991), while noting the homology between humans and nonhuman primates in the organization of cortical visual systems into "what" and "where" processing streams, also note some displacement, in humans, in the location of these systems due to development of phylogenetically newer cortical areas. They speculate that this may have ramifications for "functions that humans do not share with nonhuman primates, such as language." Similar homology among humans and nonhuman primates, with some displacement of areas specialized for spatial working memory in humans, is noted by Ungerleider et al. (1998), who also speculate that this displacement is related to the emergence of distinctively human cognitive abilities.

The broad separation of visual pathways into ventral and dorsal has been tested against performance on a range of spatial tasks in normal individuals (Chen et al. 2000). Seven spatial tasks were administered, of which three "were constructed so as to rely primarily on known ventral stream functions and four were constructed so as to rely primarily on known dorsal stream functions" (p. 380). For example, a task where subjects had to make a same/different judgement on pairs of random irregular shapes was classified as a task depending largely on the ventral stream; and a task in which "participants had to decide whether two buildings in the top view were in the same locations as two buildings in the side view" (p. 383) was classified as depending largely on the dorsal stream. These classifications, though subtle, seem consistent with the general tenor of the research reviewed here, namely, that recognition of the properties of objects is carried out via the ventral stream and the spatial location of objects is carried out via the dorsal stream. After statistical analysis of the performance of forty-eight subjects on all these tasks, Chen et al. conclude

> The specialization for related functions seen within the ventral stream and within the dorsal stream have direct behavioral manifestations in normal individuals. . . . at least two brain-based ability factors, corresponding to the functions of the two processing streams, underlie individual differences in visuospatial information processing. (Chen et al. 2000, p. 386)

Chen et al. speculate that the individual differences in ventral and dorsal abilities have a genetic basis, mentioning interesting links with Williams syndrome (Bellugi et al. 1988; Frangiskakis et al. 1996).

Milner (1998) gives a brief but comprehensive overview of the evidence, up to 1998, for separate dorsal and ventral streams in vision. For my purposes, Pylyshyn (2000) sums it up best:

> The most primitive contact that the visual system makes with the world (the contact that precedes the encoding of any sensory properties) is a contact with what have been termed visual objects or proto-objects . . . As a result of the deployment of focal attention, it becomes possible to encode the various properties of the visual objects, including their location, color, shape and so on. (Pylyshyn 2000, p. 206)

2.1.2. Auditory location and recognition. Less research has been done on auditory systems than on vision. There are recent indications that a dissociation exists between the spatial location of the source of sounds and recognition of sounds, and that these different functions are served by separate neural pathways.

Rauschecker (1997), Korte and Rauschecker (1993), and Tian and Rauschecker (1998) investigated the responses of single neurons in cats to various auditory stimuli. Rauschecker concludes

> The proportion of spatially tuned neurons in the AE [= anterior ectosylvian] and their sharpness of tuning depends on the sensory experience of the animal. This and the high incidence of spatially tuned neurons in AE suggests that the anterior areas could be part of a "where" system in audition, which signals the location of sound. By contrast, the posterior areas of cat auditory cortex could be part of a "what" system, which analyses what kind of sound is present. (Rauschecker 1997, p. 35)

Rauschecker suggests that there could be a similar functional separation in monkey auditory cortex.

Romanski et al. (1999) have considerably extended these results in a study on macaques using anatomical tracing of pathways combined with microelectrode recording. Their study reveals a complex network of connections in the auditory system (conveniently summarized in a diagram by Kaas & Hackett 1999). Within this complex network it is possible to discern two broad pathways, with much cross talk between them but nevertheless somewhat specialized for separate sound localization and higher auditory processing, respectively. The sound localization pathway involves some of the same areas that are centrally involved in visual localization of stimuli, namely, dorsolateral prefrontal cortex and posterior parietal cortex. Kaas and Hackett (1999), in their commentary, emphasize the similarities between visual, auditory, and somatosensory systems, each dividing along "what" versus "where" lines.[3] Graziano et al. (1999) have shown that certain neurons in macaques have spatial receptive fields limited to about 30 cm around the head of the animal, thus contributing to a specialized sound-location system.

Coming to human audition, Clarke et al. (2000) tested a range of abilities in four patients with known lesions, concluding

> Our observation of a double dissociation between auditory recognition and localisation is compatible with the existence of two anatomically distinct processing pathways for non-verbal auditory information. We propose that one pathway is involved in auditory recognition and comprises lateral auditory areas and the temporal convexity. The other pathway is involved in auditory-spatial analysis and comprises posterior auditory areas, the insula and the parietal convexity. (Clarke et al. 2000, p. 805)

Evidence from audition is less central to my argument than evidence from vision. My main claim is that in predicate-argument structure, the predicate represents some judgement about the argument, which is canonically an attended-to object. There is a key difference between vision and hearing. What is seen is an *object,* typically enduring; what is heard is an *event,* typically fleeting. If language is any guide (which it surely is, at least approximately) mental sound predicates can be broadly subdivided into those which simply classify the sound itself (rendered in English with such words as *bang, rumble, rush*), and those which also classify the event or agent which caused the sound (expressed in English by such words as *scrape, grind, whisper, moan, knock, tap*). (Perhaps this broad dichotomy is more of a continuum.) When one hears a sound of the first type, such as a bang, there is no object, in the ordinary sense of "object," which "is the bang." A bang is an ephemeral event. One cannot attend to an isolated bang in the way in which one directs one's visual attention to an enduring object. The only way one can simulate attention to an isolated bang is by trying to hold it in memory for as long as possible. This is quite different from maintained visual attention, which gives time for the ventral stream to do heavy work categorizing the visual stimuli in terms of complex properties. Not all sounds are instantaneous, like bangs. One can notice a continuous rushing sound. But again, a rushing sound is not an object. Logically, it seems appropriate to treat bangs and rushing sounds either with zero-place predicates, that is, as predicates without arguments, or as predicates taking event variables as arguments. (The exploration of event-based logics is a relatively recent development.) English descriptions such as *There was a bang* or *There was a rushing* tend to confirm this.

Sounds of the second type, classified in part by what (probably) caused them, allow the hearer to postulate the existence of an object to which some predicate applies. If, for example, you hear a miaow, you mentally classify this sound as a miaow. This, as with the bang or the rushing sound, is the evocation of a zero-place predicate (or alternatively a predicate taking an event variable as argument). Certainly, hearing a miaow justifies you in inferring that there is an object nearby satisfying certain predicates, in particular *CAT(x).* But it is vital to note that the English word *miaow* is two-ways ambiguous. Compare *That sound was a miaow* with *A cat miaowed,* and note that you can't say **That sound miaowed* or **That cat was a miaow.* Where the subject of *miaow* describes some animate agent, the verb actually means "cause a miaow sound."

It is certainly interesting that the auditory system also separates "where" and "what" streams. But the facts of audition do not fit so closely with the intuitions, canonically involving categorizable enduring objects, which I believe gave rise to the invention by logicians of predicate-argument notation. The idea of zero-place predicates has generally been sidelined in logic (despite their obvious applicability to weather phenomena); and the extension of predicate-argument notation to include event variables is relatively recent. (A few visual predicates, like that expressed by English *flash,* are more like sounds, but these are highly atypical of visual predicates.)

We have now considered both visual and auditory perception, and related them to object-location motor responses involving eye movement, head movement, body movement, and manual grasping. Given that when the head moves, the eyes move too, and when the body moves, the hands, head, and eyes also move, we should perhaps not be surprised to learn that the brain has ways of controlling the interactions of these body parts and integrating signals from them into single coherent overall responses to the location of objects. Given a stimulus somewhere far round to one side, we instinctively turn our whole body toward it; if the stimulus comes from not very far around, we may only turn our head; and if the stimulus comes from quite close to our front, we may only move our eyes. All this happens regardless of whether the stimulus was a heard sound or something glimpsed with the eye. Furthermore, as we turn our head or our eyes, light from the same object falls on a track across the retina, yet we do not perceive this as movement of the object. Research is beginning to close in on the areas of the brain that are responsible for this integrated location ability. Duhamel et al. (1992) found that the receptive fields of neurons in lateral intraparietal cortex are adjusted to compensate for saccades.

> One important form of spatial recoding would be to modulate the retinal information as a function of eye position with respect to the head, thus allowing the computation of location in head-based rather than retina-based coordinates. . . . By the time visual information about spatial location reaches premotor areas in the frontal lobe, it has been considerably recalibrated by information derived from eye position and other non-retinal sources. (Milner & Goodale 1995, p. 90)

The evidence that Milner and Goodale (1995) cite is from Galletti and Battaglini (1989), Andersen et al. (1985; 1990), and Gentilucci et al. (1983). Brotchie et al. (1995) present evidence that in monkeys

> the visual and saccadic activities of parietal neurons are strongly affected by head position. The eye and head position effects are equivalent for individual neurons, indicating that the modulation is a function of gaze direction, regardless of whether the eyes or head are used to direct gaze. These data are consistent with the idea that the posterior parietal cortex contains a distributed representation of space in body-centred coordinates. (Brotchie et al. 1995, p. 232)

Gaymard et al. (2000, p. 819) report on a pathological human case which "supports the hypothesis of a common unique gaze motor command in which eye and head movements would be rapidly exchangeable." Nakamura (1999) gives a brief review of this idea of integrated spatial representations distributed over parietal cortex. Parietal cortex is the endpoint of the dorsal stream, and neurons in this area both respond to visual stimuli and provide motor control of grasping movements (Jeannerod et al. 1995). In a study of vision-guided manual reaching, Carrozzo et al. (1999) have located a gradual transformation from viewer-centered to body-centered and arm-centered coordinates in superior and inferior parietal cortex. Graziano et al. (1997) discovered "arm+visual" neurons in macaques, which are sensitive to both visual and tactile stimuli, and in which the visual receptive field is adjusted according to the position of the arm. Stricanne et al. (1996, p. 2071) investigated how lateral intraparietal (LIP) neurons respond when a monkey makes saccades to the remembered location of sound sources in the absence of visual stimulation; they propose that "area LIP is either at the origin of, or participates in, the transformation of auditory signals for oculomotor purposes." Most recently, Kikuchi-Yorioka and Sawaguchi (2000) have found neurons which are active both in the brief remembering of the location of a sound and in the brief remembering of the location of a light stimulus. A further interesting connection between visual and auditory localization comes from Weeks et al. (2000), who find that both sighted and congenitally blind subjects use posterior parietal areas in localizing the source of sounds, but the blind subjects also use right occipital association areas originally intended for dorsal-stream visual processing. Egly et al. (1994) found a difference between left-parietal-lesioned and right-parietal-lesioned patients in an attention-shifting task.

The broad generalization holds that the dorsal stream provides very little of all the information about an object that the brain eventually gets, but just about enough to direct attention to its location and enable some motor responses to it. The ventral stream fills out the picture with further detailed information, enough to enable a judgement by the animal about exactly what kind of object it is dealing with (e.g., flea, hair, piece of grit, small leaf, shadow, nipple; or, in another kind of situation, brother, sister, father, enemy, leopard, human). A PET scan study (Martin et al. 1996) confirms that the recognition of an object (say, as a gorilla or a pair of scissors) involves activation of a ventral occipitotemporal stream. The particular properties that an animal identifies will depend on its ecological niche and lifestyle. It probably has no need of a taxonomy of pieces of grit, but it does need taxonomies of fruit and prey animals, and will accordingly have somewhat finely detailed mental categories for different types of fruit and prey. I identify such mental categories, along with non-constant properties, such as colour, texture, and movement, which the ventral stream also delivers, with *predicates.*

2.2. "Dumb" attentional mechanisms and the object/property distinction

Some information about an object, for example, enough about its shape and size to grasp it, can be accessed via the dorsal stream, in a preattentive process. The evidence cited above from optical size illusions in normal subjects shows that information about size as delivered by the dorsal stream can be at odds with information about size as delivered by the ventral stream. Thus, we cannot say that the two streams have access to exactly the same property, "size"; presumably the same is true for shape. Much processing for shape occurs in the ventral stream, after its divergence from the dorsal stream in V1 (Gross 1992); at the early V1 stage, full shapes are not represented, but rather basic information about lines and oriented edges, as Hubel and Wiesel (1968) first argued, or possibly about certain 3D aspects of shape (Lehky & Sejnowski 1988). Something about the appearance of an object in peripheral vision draws attention to it. Once the object is focally attended to, we can try to report the "something" about it that drew our attention. But the informational encapsulation (in the sense of Fodor 1983) of the attention-directing reflex means that the more deliberative process of contemplating an object cannot be guaranteed to report accurately on this "something." And stimuli impinging on the retinal periphery trigger different processes from stimuli impinging on the fovea. Thus, it is not clear whether the dorsal stream can be said to deliver any properties, or mental predicates, at all. It may not be appropriate to speak of the dorsal stream delivering *representations* (accessible to report) of the nature of objects. Nevertheless, in a clear sense, the dorsal stream does *deliver objects,* in a minimal sense of "object" to be discussed below. What the dorsal stream delivers, very fast, is information about the egocentric location of an object, which triggers motor responses resulting in the orientation of focal attention to the object. (At a broad-brush level, the differences between preattentive processes and focal attention have been known for some time, and are concisely and elegantly set out in Ch. 5 of Neisser 1967.)

In a functioning high-level organism, the information provided by the dorsal and ventral streams can be expected to be well coordinated (except in the unusual circumstances which generate illusions). Thus, although predicates/properties are delivered by the ventral stream, it would not be surprising if a few of the mental predicates available to a human being did not also correspond at least roughly to information of the type used by the dorsal stream. But humans have an enormous wealth of other predicates as well, undoubtedly accessed exclusively via the ventral stream, and bearing only indirect relationships to salient attention-drawing traits of objects. Humans classify and name objects (and substances) on the basis of properties at all levels of concreteness and salience. Landau et al. (1988; 1998a; 1998b) and Smith et al. (1996) report a number of experiments on adults' and children's dispositions to name familiar and unfamiliar objects. There are clear differences between children and adults, and between children's responses to objects that they in some sense understand and to those that are strange to them. Those subjects with least conceptual knowledge of the objects presented, that is, the youngest children, presented with strange objects, tended to name objects on the basis of their shape. Smith et al. (1996) relate this disposition to the attention-drawing traits of objects:

> Given that an adult is attending to a concrete object and producing a novel name, children may interpret the novel name as referring to "whatever it is about the object that most demands attention." An attentional device that produces this result may work well enough to *start* a child's learning of a specific object name. (Smith et al. 1996, p. 169)

This is not unexpected. Higher-level features and categories are learned, and once learned, can be applied in extending names to things. The youngest humans, having learned few or no higher-level categories, have only the most basic features to appeal to, those corresponding to information gleaned by the dorsal stream. See Bloom (2000) for a recent commentary on this literature, emphasizing a different theme, but consistent with the hypothesis that children's *earliest* naming tendencies capitalize strongly on attention-drawing traits of objects.

But doesn't talk of "attention-drawing traits of objects" undermine my central argument, by locating some "traits" (alias properties) within the class of information delivered by the dorsal stream? A position diametrically opposed to mine would be that ultimately there is no distinction at all to be made between objects and properties. A philosophical argument for such a position might appeal to English terms such as "objecthood," meaning the property of being an object. Advanced logical systems can play havoc with basic ontological categories, such as object and property, by various devices such as type-raising. Such devices may be appropriate in the analysis of elaborated human languages and the systems of thought that they make available. Yes, humans can treat properties as objects, by reification, and objects as properties (by "Pegasizing Pegasus," as Quine put it). But I would claim that an ape's mental traffic with the world is in terms of two broadly noninterconvertible ontological categories, object and property.

A more psychologically plausible argument against my position might claim that *any* property of an object that one could give a name to could in principle be an attention-drawing trait. This would potentially attribute to the dorsal stream any information conveyed by a predicate, thus destroying the hypothesis that it is the ventral stream that delivers predicates. I emphasize that such issues should be addressed with empirical (neuro-) psychological evidence, rather than purely philosophical argumentation. Some relevant evidence exists, pointed out by O'Brien and Opie (1999), in connection with blindsight, as follows.

> Consider the comments made by Weiskrantz' subject D.B., after performing well above chance in a test that involved distinguishing between Xs and Os presented in his scotoma. While D.B. maintained that he performed the task merely by guessing:
>
>> If pressed, he might say that he perhaps had a "feeling" that the stimulus was either pointing this or that way, or was "smooth" (the O) or "jagged" (the X). On one occasion in which "blanks" were randomly inserted in a series of stimuli . . . he afterwards spontaneously commented he had a feeling that maybe there was no stimulus present on some trials. But always he was at a loss for words to describe any conscious perception, and repeatedly stressed that he saw nothing at all in the sense of "seeing," and that he was merely guessing. (Weiskrantz et al. 1974, p. 721)
>
> Throughout D.B.'s verbal commentaries there are similar remarks. Although he steadfastly denies "seeing" in the usual way when presented with visual stimuli, he frequently describes some kind of concurrent awareness. He talks of things "popping out a couple of inches" and of "moving waves," in response to single point stimuli (Weiskrantz 1986, p. 45). He also refers to "kinds of pulsation" and of "feeling some movement" in response to moving line stimuli. (Weiskrantz 1986, p. 67)
>
> Consequently, while blindsight subjects clearly do not have normal visual experience in the "blind" regions of their visual fields, this is not to say that they don't have any phenomenal experience whatsoever associated with stimuli presented in these regions. What is more, it is not unreasonable to suggest that what little experience they do have in this regard explains their residual discriminative abilities. D.B., for example, does not see Xs or Os (in the conventional sense). But in order to perform this task he doesn't need to. All he requires is some way of discriminating between the two stimulus conditions *q* some broad phenomenal criterion to distinguish "Xness" from "Oness." And as we've seen, he does possess such a criterion: one stimulus condition feels "jagged" while the other feels "smooth." Thus, it is natural to suppose that he is able to perform as well as he does (above chance) because of the (limited) amount of information that is consciously available to him. (O'Brien & Opie 1999, p. 131)

Unlike O'Brien and Opie, I am not mainly concerned with consciousness. I am content to concede that O'Brien and Opie have a point, and to fall back on the reservation that a formula as simple as *PREDICATE(x)* cannot be expected to mirror exactly all the processes of such a complex organ as the brain. The stark contrast between the blindsight patient's experience and his performance is evidence that the brain separates sub- or semiconscious awareness of the bare presence of an object from the vast array of judgements that can be made by a normal person about the properties of an object. Perhaps training can boost the set of properties which can act as attention-drawing traits. But I would predict that only a tiny subset of properties are natural attention-drawing properties, and that any properties added to this set by practice or training are likely to swing into action significantly more slowly than the primal attention-drawing properties.

This prediction conflicts with a prediction of Milner and Goodale's in their final chapter addressing further research questions prompted by the dorsal/ventral distinction. They write "It is unlikely that the dorsal stream plays the major role in mediating this initial [attention] selection process, since object recognition and 'semantic' knowledge may have to be taken into account." (Milner & Goodale 1995, p. 202) With due deference to Milner and Goodale, I suggest that their implicit premise that *all* "semantic" recognition takes place in the ventral stream may be too strong, and that a very limited set of primal properties can be accessed by the dorsal stream. I would further claim that access to these primal attention-drawing properties is highly encapsulated, unlike access to properties delivered by the ventral stream. It is an intuition of this difference that gives rise to the logician's postulate that the fundamental logical structure is an asymmetric relation between two distinct logical types, predicate and argument.

As an interim summary, the formula *PREDICATE(x)* is a simplifying schematic representation of the integration by the brain of two broadly separable processes. One process is the rapid delivery by the senses (visual and/or auditory) of information about the egocentric spatial location of a referent object relative to the body, represented in parietal cortex. The eyes, often the head and body, and sometimes also the hands, are oriented to the referent object, which becomes the instantiation of a mental variable. The other process is the somewhat slower analysis of the delivered referent object by the perceptual (visual or auditory) recognition subsystems in terms of its properties. The asymmetric relationship between the predicate and the variable, inherent in the bracketing of the formula, also holds of the two neural processes:

> From the genetical and functional perspectives, the two modes of processing are asymmetrically related: while egocentric evaluation of "where" need not take into account the identity of objects, the perception of "what" usually proceeds through an intermediate stage in which objects are dynamically localized. (Bridgeman et al. 1994)

There is an interesting parallel (more than merely coincidental) in the uses of the term "binding" in logic and neuroscience. The existence of a blue dot can be represented in FOPL as $\exists x$ *[BLUE(x) & DOT(x)]*. (The ordering of the conjuncts is immaterial.) Here the existential quantifier is said to "bind" the variable *x* immediately after it, and, also importantly, all further instances of this variable must fall within the scope, indicated by brackets, of the quantifier. The variable and its binding quantifier thus serve to unite the various predicates in the formula, indicating that they denote properties of the *same* object. Logical binding is not a relationship between a predicate and its argument, but a relationship between all predicates in the scope of a particular quantifier which take the bound variable as argument. In neuroscience, "*binding* is the problem of representing conjunctions of properties. . . . For example, to visually detect a vertical red line among vertical blue lines and diagonal red lines, one must visually bind each line's color to its orientation" (Hummel 1999). Detection of properties is generally achieved via the ventral stream. The dorsal stream directs attention to an object. Once attention is focussed on a particular object, the ventral stream can deliver a multitude of different judgements about it, which can be represented logically by a conjunction of 1-place predications. The bare drawing of attention to an object, with no category judgements (yet) made about it, corresponds to the "$\exists x$" part of the logical formula.

Evidently, the brain does solve the binding problem, although we are not yet certain exactly how it does it. The claim advanced here for a connection between predicate-argument structure and the ventral/dorsal separation does not depend on what, in detail, the brain's solution to the binding problem turns out to be.

2.3. Related proposals

2.3.1. Landau and Jackendoff: Nouns and prepositions. Jackendoff and Landau (1992) and Landau and Jackendoff (1993)[4] noticed the early neurological literature on ventral and dorsal streams, and proposed a connection between the "where"/"what" dichotomy and the linguistic distinction between prepositions and common nouns. They correlate common nouns denoting classes of physical objects with information provided by the ventral stream, and prepositions with information provided by the dorsal stream. Landau and Jackendoff emphasize the tentative and suggestive nature of their conclusions, but it will be useful to explain briefly why I believe their proposed correlations are incorrect, and to contrast their suggestions with mine.

Let us start with the proposed noun/ventral correlation. Nouns, as they correctly state, encode complex properties, such as being a dog. And categorization of objects, as when one recognizes a particular object as a dog, involves the ventral stream. This much is right. Landau and Jackendoff emphasize the striking contrast between the enormous number of nouns in a language and the very restricted number of prepositions. It is this stark quantitative contrast which stands in need of explanation, and for which they invoke the neurological "what"/"where" distinction. Their reasoning is that the dorsal stream provides a bare minimum of information about the location of an object (no more than is encoded by the small inventory of prepositions in a language), while the ventral stream does all the rest of the work that may be necessary in categorizing it. This characterization of the relative amounts of linguistically expressible information provided by the respective streams certainly goes in the right direction (but is in fact, I will argue, an understatement).

However, a correlation of populous syntactic categories (such as noun) with the ventral stream, and a complementary correlation of sparsely populated categories (such as preposition) with the dorsal stream will not work. Consider adjectives. Adjectives are never as numerous in a language as nouns, many languages have only about a dozen adjectives, and some languages have none at all (Dixon 1982). Taking the numbers of nouns, adjectives, and prepositions (or postpositions) across languages as a whole, one would be more likely to group adjectives with prepositions as relatively sparsely populated syntactic categories. But many of the properties typically expressed by adjectives, such as colour, are detected within the ventral stream. Landau and Jackendoff might respond with the revised suggestion that the ventral stream processes both noun meanings and adjective meanings, leaving the difference in typical numbers of nouns and adjectives still unexplained, and this is fair enough, but it gets closer to the correlation proposed in the present paper between predicates generally and the ventral stream. Indeed, when one considers all syntactic categories, rather than restricting discussion to just nouns and prepositions, it is clear that judgements corresponding to the meanings of many verbs (e.g., *move* and its hyponyms), and many adverbs (e.g., *fast* and similar words) are made in the ventral stream. Verbs are pretty numerous in languages, though not as numerous as nouns, while adverbs are much less numerous, and some languages don't have adverbs at all. The relative population-size of syntactic categories does not correlate with the ventral/dorsal distinction.

Now consider Landau and Jackendoff's proposed dorsal/preposition correlation. Prepositions express predicates, many of which give spatial information, both egocentric and allocentric. Their article naturally depended on the literature available at the time it was written, especially the classic Ungerleider and Mishkin (1982), which gave the impression of a distinction between "object vision" and a single system of "spatial vision." In a later very detailed critique of this work, Milner and Goodale (1995) devote several chapters to accumulating evidence that an egocentric system of "visual guidance of gaze, hand, arm or whole body movement" (p. 118) is located in the posterior parietal region, while many other kinds of visual judgement, including computation of allocentric spatial information, are made using occipito-temporal and infero-temporal regions of cortex. "Perhaps the most basic distinction that needs to be made in thinking about spatial vision is between the locational coordinates of some object within the visual field and the relationship between the loci of more than one object" (Milner & Goodale 1995, p. 89). Prepositions do not respect this distinction, being used indiscriminately for both egocentric (e.g., *behind me*) and allocentric (e.g., *behind the house*) information. Only information of the egocentric kind is computed in the dorsal stream.

Of course, as Bryant (1993, p. 242) points out, there must be interaction between the systems for egocentric location and the building of allocentric spatial maps. Galati et al. (2000) is a recent fMRI study which begins to relate egocentric and allocentric functions to specific regions of cortex.

Both nouns and prepositions express predicates. I have argued that the categorical judgements of properties and relations involved in the application of all predicates to attended-to objects are mediated by the ventral stream. The key logical distinction is between predicates and individual variables, not between different syntactic subclasses of words which express predicates. Thus, the logical distinction correlated here with the neurological dorsal/ventral distinction is considerably more fundamental, and hence likely to be evolutionarily more primitive, than the distinction on which Landau and Jackendoff focus. This idea is close to what I believe Bridgeman (1993), in his commentary on them, states: "cognitive and [motor-oriented] spatial systems can be distinguished on a lower level than that of Landau and Jackendoff, a level that differentiates linguistic from nonlinguistic coding" (p. 240). Predicates are coded linguistically; the vast majority of words in a language correspond to predicates. In languages generally, only a tiny inventory of words, the indefinite pronouns, such as *something* and *anything* could be said to correlate directly with the individual variables x, y, z of simple formulae such as $\exists x\ [LION(x)]$, loosely translatable as *Something is a lion.* In more complex examples, a case can be made that the logical variables correspond to anaphoric pronouns, as in *There was a lion and it* yawned. The deictic nature of the variables whose instantiations are delivered to posterior parietal cortex by the sensory "where" systems will be the subject of section 4.

2.3.2. Givon: Lexical concepts and propositions. Givon (1995, pp. 408–10), in a brief but pioneering discussion, relates the dorsal and ventral visual pathways to linguistic information in a way which is partly similar to my proposed correlation. In particular, Givon correlates information accessed via the ventral stream with lexical concepts. This is very close to my correlation of this information with prelinguistic predicates. Prelinguistic predicates are concepts (or what Bickerton calls "protoconcepts"), and they can become *lexical* concepts by association with phonological forms, once language gets established. My proposal differs from Givon's in the information that we correlate with the dorsal stream, which he correlates with "spatial relation/motion – propositional information about states or events" (p. 409). Givon, writing before 1995, relied on several of the same sources as Landau and Jackendoff, and, like them, assumes that "the dorsal (upper) visual processing stream analyzed the *spatial relations* between specific objects and *spatial motion* of specific objects. This processing track is thus responsible for analyzing specific visual *states* and *events*" (p. 409, emphasis in original). As mentioned above, Milner and Goodale (1995) subsequently presented evidence that such allocentric spatial information is not processed in the dorsal stream. Elsewhere in Givon's account, there is an acknowledgement of the role of the stream to the temporal lobe in accessing information about spatial motion:

> Further, even in non-human primates, the object recognition (ventral) stream analyzes more than visually perceived objects and their attributes. Thus Perrett et al. (1989) in their study of single-cell activation in monkeys have been able to differentiate between single cortical cells that respond to objects (nouns), and those that are activated by *actions* (verbs). Such differentiation occurs within the object recognition stream itself, in the *superior temporal sulcus* of the left-temporal lobe. And while the verbs involved – e.g., moving an object by hand towards mouth – are concrete and spatio-visual, they involve more abstract computations of purpose and causation. (Givon 1995, p. 410, emphasis in original)

This attribution undermines Givon's earlier identification of the dorsal stream as the stream providing information about spatial motion. Note that Givon begins to correlate neural structure with the specifically linguistic categories of noun and verb, a move which I avoid. I correlate information accessed by the ventral stream with predicates, regardless of whether these eventually get expressed as nouns, verbs, adjectives, or any other lexical category. The present proposed correlation of distinct neural pathways with *logical* predicates and individual variables differs from both Landau and Jackendoff's and Givon's proposals in claiming completely *prelinguistic* correlates for the ventral and dorsal pathways. The correlation that I propose for information delivered by the dorsal stream is developed in more detail in the next section.

2.3.3. Rizzolatti and Arbib: A prelinguistic "grammar" of action. Rizzolatti and Arbib's paper (1998) contains a section entitled "A pre-linguistic 'grammar' of action in the monkey brain." Like me, they are concerned with a neural precursor to language, found in monkey brains. There are superficial similarities between our proposals and differences, which are important to state.

Rizzolatti and Arbib use a kind of logical notation to convey an idea about the activity of "canonical" macaque F5 neurons in grasping small objects.

> We view the activity of "canonical" F5 neurons as part of the code for an imperative case structure, for example, Command: grasp-A(raisin), as an instance of grasp-A(object), where grasp-A is a specific kind of grasp, to be applied to the raisin. Note that this case structure is an "action description," not a linguistic representation. "Raisin" denotes the specific object towards which the grasp is directed, whereas grasp-A is a specific command directed towards an object with well specified physical properties. (Rizzolatti & Arbib 1998, p. 192)

The formula used here by Rizzolatti and Arbib is best taken as a shorthand for a sequence of separate processes; the compression into a single formula gives rise to several potentially misleading infelicities. Logically, a term like "raisin" is a predicate, and therefore (in FOPL) should not be used as an argument. This is not a merely pernickety point. Key to my own proposal is the idea that a *predicate* is the logical expression of a judgement about the category to which some attended-to object belongs. The process of perceiving something to be a raisin is, I claim, well represented by the formula *RAISIN(x).* Allowing, for the moment, "*GRASP-A*" as a predicate, the sequence of events in the monkey's brain with which R&A are here concerned would be better expressed as

RAISIN(x)	Equation 1
GRASP-A(x)	Equation 2

That is, the judgement that the attended-to object is a raisin precedes the motor instruction to grasp it in a certain way, if the animal is acting with any deliberation. If the animal does not make a deliberate categorical judgement, but simply reflexively grabs the object (with activation essentially limited to the dorsal stream), then, according to the correlation I propose, there is no question of the predicate *RAISIN*, or any other predicate, being involved. I have less to say about the use of predicate notation to cover motor instructions. Classical logic was devised as a way of objectively representing (inter alia) observable events and states of affairs, and the present proposal is to link logic to the neural processes involved in *perception* of stimuli from outside the animal, and not to the mechanisms involved in purposeful action by the animal. Rizzolatti and Arbib's discussion, while appealing to a notation which is logic-like in that it apparently has predicate-argument structure, does not in fact deconstruct this formula and attribute the separate parts to different neural processes, as is proposed in the present article.

3. Attention to locations, features, or objects?

Thus far, I have correlated logical predicates with perceived features, such as colour or shape, or more complex combinations of features, which make up a particular face; and I have correlated the instantiations of individual variable arguments of predicates with whole objects attended to, such as a particular bird, stone or tree. But, one might ask, isn't an object nothing more than a bundle of features?[5] The notion of an object, as opposed to its features, is important for the central claim of this article, that modern neuroscience has revealed close correlates of the elements of the logical *PREDICATE(x)* formula. In FOPL, individual variables are instantiated by whole objects, not by properties. Substantial evidence now exists that the primary targets of attentive processes are indeed whole objects, and not properties or features.

Beside the object/feature distinction, the object/location distinction must also be mentioned. Preattentive processes, operating largely through the dorsal stream, direct attention to a location represented in a mental spatial map defined in terms of parts of the body. So, in a sense, attention is directed to a place, rather than to an object. But, except in cases of illusion or stimuli that vanish as soon as they are noticed, what the mind finds at the location to which attention is directed is an object. So what is held in attention, the object, or the location? Evidence has accumulated in recent years that what is held in attention are objects, and not locations.

A paper by Duncan (1984), while by no means the first on this topic, is a good place to start a survey of recent research. Duncan distinguishes between object-based, discrimination-based, and space-based theories of visual attention.

> *Object-based* theories propose a limit on the number of separate objects that can be perceived simultaneously. *Discrimination-based* theories propose a limit on the number of separate discriminations that can be made. *Space-based* theories propose a limit on the spatial area from which information can be taken up. (p. 501)

Space-based theories have been called "mental spotlight" theories, as they emphasize the "illumination" of a small circle in space. Duncan experimented with brief exposures to narrow displays, subtending less than one degree at the eye, consisting of two overlapping objects, an upright box (small or large) with a line (dotted or dashed) passing down through it. The box always had a small gap in one side, to left or right, and the line always slanted slightly to the right or the left. Subjects had to report judgements on two dimensions at a time, from the four possible dimensions *box(size)*, *box(gap)*, *line(tilt)*, and *line(texture)*.

> It was found that two judgments that concern the same object can be made simultaneously without loss of accuracy, whereas two judgments that concern different objects cannot. Neither the similarity nor the difficulty of required discriminations, nor the spatial distribution of information, could account for the results. The experiments support a view in which parallel, preattentive processes serve to segment the field into separate objects, followed by a process of focal attention that deals with only one object at a time. (p. 501)

And,

> The present data confirm that focal attention acts on packages of information defined preattentively and that these packages seem to correspond, at least to a first approximation, to our intuitions concerning discrete objects. (Duncan 1984, p. 514)

Duncan notes that object-based, discrimination-based, and space-based theories are not mutually exclusive. This idea is repeated by some later writers (e.g., Egly et al. 1994; Vecera & Farah 1994), who discuss the possibilities of distinct systems of attention operating at different stages or levels (e.g., early versus late) or in response to different tasks (e.g., expectancy tasks versus selection tasks). The experimental evidence for space-based attention provided by these authors involves a different task from the task that Duncan set his subjects (although the experimental materials were very similar). Duncan asked his subjects for judgements about the objects attended to. The experiments suggesting space-based attention involved subjects being given a "precue" (mostly valid, sometimes not) leading them to expect a stimulus to appear in a certain area, or on a certain object, and their task was simply to press a button when the stimulus appeared. Reaction times were measured and compared. Vecera and Farah (1994) suggest: "Instead of attention being a single limitation or a single system, there may be different types of limitations or different types of attention that depend on the representations used in different tasks" (p. 153). This way of expressing it seems to me to depart from the useful distinction between *preattentive* processes and *focal attention.* Duncan's subjects gave judgements about what was in their focal attention. In the precued experiments, the reaction times measured the subjects' preattentive processes. As Egly et al. (1994) note, "previous findings revealed evidence for both space-based and object-based components to visual attention. However, we note that these two components have been identified in very different paradigms" (p. 173). I will continue on the assumption that the cued reaction-time paradigm in fact tests preattentive processes. My question here is whether focal attention operates on objects, locations, or features.[6]

A series of papers (Baylis 1994; Baylis & Driver 1993; Gibson 1994) takes up Duncan's theme of whether focal attention is applied to objects or locations. As with Duncan's experiments, subjects were required to make judgements about what they saw, but in this case reaction times were measured. In most of the experiments, the displays shown to subjects could be interpreted as either a convex white object against a black ground, or two partly concave black objects with a white space between them. Subjects had to judge which of two apices in the display was the lower. The apices could be seen as belonging to the same (middle) object, or to two different (flanking) objects.

> Position judgments about parts of one object were more rapid than equivalent judgments about two objects even though the positions to be compared were the same for one- and two-object displays. This two-object cost was found in each of five experiments. Moreover, this effect was even found when the one- and two-object displays were physically identical in every respect but parsed as one or two objects according to the subjects' perceptual set. . . . We propose that spatial information is routinely represented in two different ways in the visual system. First, a scene-based description of space represents the location of objects within a scene. Second, an object-based description is produced to describe the relative positions of parts of each object. Such a hierarchical representation of space may parallel the division of the primate visual system into a scene-based dorsal stream and an object-based ventral stream." (Baylis & Driver 1993, pp. 466–67)

Gibson (1994) suggested that these results could have been caused by a confound between the number of objects perceived and the concavity or convexity of the objects. Baylis (1994) replied to this objection with further experiments controlling against this possible confound, reinforcing the original conclusion that making a judgement about two objects is more costly than making a judgement about a single object, even when the displays are in fact physically identical.

Luck and Vogel (1997) presented subjects with visual arrays, with a slight delay between them, and asked them to report differences between the arrays. They summarize their conclusion as follows:

> It is possible to retain information about only four colours or orientations in visual working memory at one time. However, it is also possible to retain both the colour and the orientation of four objects, indicating that visual working memory stores integrated objects rather than individual features. Indeed, objects defined by a conjunction of four features can be retained in working memory just as well as single-feature objects, allowing sixteen individual features to be retained when distributed across four objects. Thus, the capacity of visual working memory must be understood in terms of integrated objects rather than individual features. (p. 279)

Valdes-Sosa et al. (1998) studied transparent motion "defined by two sets of differently colored dots that were interspersed in the same region of space, and matched in spatial and spatial frequency properties" (p. B13).

> Each set moved in a distinct and randomly chosen direction. We found that simultaneous judgments of speed and direction were more accurate when they concerned only one set than when they concerned different sets. Furthermore, appraisal of the directions taken by two sets of dots is more difficult than judging direction for only one set, a difficulty that increases for briefer motion. We conclude that perceptual grouping by common fate exerted a more powerful constraint than spatial proximity, a result consistent with object-based attention. (p. B13)

The most recent and most ingenious experiment comparing object-based, feature-based, and location-based theories of attention is Blaser et al. (2000). In this experiment, subjects were presented with a display consisting of two patterned patches ("Gabors"), completely spatially superimposed. The trick of getting two objects to seem to occupy the same space at the same time was accomplished by presenting the patches in alternate video frames. The patches changed gradually, and with a certain inertia, along the three dimensions of colour, thickness of stripes, and orientation of stripes. Subjects had to indicate judgements about the movements of these patches through "feature space." In one experiment it was shown that observers are "capable of tracking a single object in spite of a spatially superimposed distractor." In a second experiment, "observers had both an instruction and a task that encouraged them to attend and track two objects simultaneously. It is clear that observers did much worse in these conditions than in the within-object conditions, where they only had to attend and track a single object."

The story so far, then, is that the brain interprets relatively abrupt discontinuities – such as change of orientation of a line, change of colour, change of brightness – together as constructing holistic visual objects which are expected to share a "common fate." It is these whole objects that are held in attention. A shift of attention from one object to another is costly, whereas a shift of attention from one feature of an object to another feature of the same object is less costly. This is consistent with the view underlying FOPL that the entities to which predicates apply are objects, and neither properties nor locations. In accepting this correlation between logic and neuropsychology we have, paradoxically, to abandon an "objective" view of objects. No perceptible physical object is ever the same from one moment of its existence to the next. Every thing changes. Objects are merely slow events. What we perceive as objects is entirely dependent on the speed our brains work at. An object is anything that naturally attracts and holds our attention. But objects are what classical logicians have had in mind as the basic entities populating their postulated universes. The tradition goes back at least to Aristotle, with his "primary substances" (= individual physical objects).

4. Computing deictic variables in vision, action, and language

The previous section concerned the holding in attention of single whole objects. We can deal with several different objects in a single task, and take in scenes containing more than one object. How do we do this, and what are the limits on the number of different objects we can manage to "keep in mind" at any one time?

The idea of objects of attention as the temporary instantiations of mental computational variables has been developed by Kahneman and Treisman (1992), Ballard et al. (1995; 1997), and Pylyshyn (2000), drawing on earlier work including Kahneman and Treisman (1984), Ullman (1984), Agre and Chapman (1987) and Pylyshyn (1989). The idea behind this work is that the mind, as a computational device for managing an organism's interactions with the world, has available for use at any time a small number of "deictic" or "indexical" variables. Pylyshyn (1989) calls such variables "FINSTs," a mnemonic for "INSTantiation FINger."

> A FINST is, in fact, a reference (or index) to a particular feature or feature cluster on the retina. However, a FINST has the following additional important property: because of the way clusters are primitively computed, a FINST keeps pointing to the "same" feature cluster as the cluster moves across the retina. . . . The FINST itself does not encode any properties of the feature in question, it merely makes it possible to locate the feature in order to examine it further if needed. (Pylyshyn 1989, pp. 69–70)

> This is precisely what the FINST hypothesis claims: it says that there is a primitive referencing mechanism for *pointing to* certain kinds of features, thereby maintaining their distinctive identity without either *recognizing* them (in the sense of categorizing them), or explicitly encoding their locations. (Pylyshyn 1989, p. 82, emphasis in original)

All practical tasks involve analysis of the scene of the task in terms of the principal objects concerned. The simple scene-descriptions of predicate logic, such as $\exists x, y\ [MAN(x)\ \&\ DOG(y)\ \&\ BEHIND(y,x)]$ (translated as *A dog is behind a man*) have direct counterparts in examples used by vision researchers of what happens in the brain when analyzing a visual scene. An early example from Ullman is:

> Suppose, for example, that a scene contains several objects, such as a man at one location, and a dog at another, and that following the visual analysis of the man figure we shift our gaze and processing focus to the dog. The visual analysis of the man figure has been summarized in the incremental representation, and this information is still available at least in part as the gaze is shifted to the dog. In addition to this information we keep a spatial map, a set of spatial pointers, which tell us that the dog is at one direction, and the man at another. Although we no longer see the man clearly, we have a clear notion of what exists where. The "what" is supplied by the incremental representations, and the "where" by the marking map. (Ullman 1984, p. 150)

Since this passage was written, in the early 1980s, vision research has substantially developed the idea of separate "where" and "what" neural pathways, dorsal and ventral respectively, as surveyed above.

The everyday tasks of primates are plausibly envisaged in such terms. Activities such as fishing for termites with a stick and eating them, or building a sleeping nest in a tree, or collaborating with others in a hunt, all involve attention to different objects while performing the task. During the task, immediate attention is shifted from one thing to another, but the small number of principal things involved in the task are not put out of mind. Crucial information about them is stored as the contents of variables, or computational pointers. The termite-fishing chimpanzee at one moment attends to the termites caught on its stick, and guides them to its mouth. Meanwhile, it still holds, as part of the ongoing larger task, information about the hole in the termite mound, though it is not visually attending to it while putting the termites in its mouth. After eating the termites, visual attention is switched back to the hole in the termite mound, and the stick is manually guided into the hole. The chimpanzee need not rediscover the properties of the hole (e.g., its size and orientation), because these properties have been stored as the contents of a computational variable.

(Managing scenes with several objects necessitates control of sameness and difference. The ape doing some practical task with several objects does not need to be able to distinguish these objects in principle from all other objects in the world, but certainly does need to distinguish among the objects themselves. This is the simple seed from which the more advanced concept of a unique-in-the-world individual may grow.)

An idea very similar to Pylyshyn's FINSTs, but slightly different in detail, is proposed by Kahneman and Treisman (1984; 1992). These authors hypothesize that the mind sets up temporary "object files" in which information about objects in a scene is stored. The object files can be updated, as the viewer tracks changes in an object's features or location. It is emphasized that the information stored in temporary object files is not the same as that which may be stored in long-term memory. But the information in object files can be matched with properties associated with objects in long-term memory, for such purposes as object recognition. When (or shortly after) objects disappear from the current scene, their object files are discarded. A file full of information is not a variable. In discussing the relationship between object files and Pylyshyn's FINSTs, Kahneman and Treisman (1992) suggest that "a FINST might be the initial phase of a simple object file before any features have been attached to it" (p. 217). This correspondence works well, apart from a reservation, which Kahneman and Treisman (1992) note, involving the possibility of there being objects with parts that are also objects. This is a detail that I will not go into here. An "empty" object file, available for information to be put into it, is computationally an uninstantiated variable, provided that it can be identified and distinguished from other such files that are also available and that may get different information put into them. The fact that object files can be updated, are temporary, and can be discarded for re-use with completely new values, underlines their status as computational variables used by the mind for the short-term grasping of scenes.

Kahneman and Treisman (1992) "assume that there is some limit to the number of object files that can be maintained at once" (p. 178). Ballard et al. (1997) stress that computational efficiency is optimized if the number of such variables is small. Luck and Vogel (1997) demonstrate a limit of four objects in visual working memory (and propose an interesting explanation in terms of the "oscillatory or temporally correlated firing patterns among the neurons that code the features of an object," p. 280). Pylyshyn assumes "a pool of four or five available indexes" (Pylyshyn 2000, p. 201). It is perhaps at first helpful to concretize these ideas by identifying the available variables in the same way as logicians do, by the letters *w, x, y,* and *z.* Neither logicians nor vision researchers wish to be tied to the claim that the mind can handle a maximum of only four variables, but hardly any examples given by them ever involve more than four separate variables. So it would seem for many practical purposes that about four variables are enough. In performing an everyday task, then, a creature such as a primate mentally juggles a parsimonious inventory of variables, *w, x, y, z,* . . . . Cowan (2001) provides a very thorough and extensive survey of studies of short term memory, concluding that there is a remarkable degree of similarity in the capacity limit in working memory observed with a wide range of procedures. A restricted set of conditions is necessary to observe this limit. It can be observed only with procedures that allow assumptions about what the independent chunks are, and the limit the recursive use of the limited-capacity store . . . The preponderance of evidence from procedures fitting these conditions strongly suggests a mean memory capacity in adults of 3 to 5 chunks, whereas individual scores appear to range more widely from about 2 up to about 6 chunks. The evidence for this pure capacity limit is considerably more extensive than that for the somewhat higher limit of 7 + 2 stimuli (Cowan 2001).

This small inventory of variables can explain other known size-limitations in humans and non-human primates. The upper limit of subitizing in humans is around 4; given a quick glance at a group of objects, a human can guess accurately how many there are, without explicit counting, up to a limit of about 4 or 5 (see Antell & Keating 1983; Gelman & Gallistel 1978; Mandler & Shebo 1982; Russac 1983; Schaeffer et al. 1974; Starkey & Cooper 1980 for some relevant studies). Both Ullman (1984, p. 151) and Pylyshyn (2000, pp. 201–202) make the connection between subitizing (which Ullman calls "visual counting") and the marking or indexing of locations in a scene. Trick and Pylyshyn (1993; 1994) explain the natural limit of subitizing in terms of the number of objects that can be involved in "pre-attentive" processing in vision. Dehaene (1997), in work on the numerical competences of many species, finds a natural difference between low numerosities up to about 3 or 4, and higher ones. For details of how this natural discontinuity at around 4 in the number sequence is reflected in the numerals, adjectives, and nouns of many human languages, see Hurford (1987; 2000a).

The simple clauses of human languages are constrained to a maximum of about four or five core arguments; indeed, most clauses have fewer than this. Presumably this reflects the structure of the underlying mental propositions. Conceivably, one could analyze the content of a complex sentence, such as *The cat chased the mouse that stole the cheese that lay in the house that Jack built* as having a single predicate *CHASE-STEAL-LIE-BUILD* and five arguments (the cat, the mouse, the cheese, the house, and Jack). But it is more reasonable to suppose that the grammatical structure of such embedded natural language clauses reflects a mental structure involving a nesting of separate propositions, each with its own simple predicate expressing a relation between just two arguments (which may be shared with other predicates).[7]

Ballard et al. (1997) give grounds why the number of variables juggled in computing practical tasks must be small (typically no more than three). Of course, most sentences in human languages are not direct representations of any practical task on the part of the speaker, like "Put the stick in the hole." Humans exchange declarative information about the world for use at later times, for example, "Your mother's coming on Tuesday." But mental scene descriptions are necessary for carrying out practical tasks of the kind that primates are capable of, and therefore pre-exist language phylogenetically. It is plausible that the type of scene descriptions used by nonhuman primates would be reused for more complex cognitive, and ultimately linguistic, purposes. I suggest that the limitation of elementary propositions to no more than about three arguments, and the typical use of even fewer arguments, derives from the considerations of computational efficiency advanced by Ballard et al. (1997).[8]

The marking, or indexing, of spatial locations in a visually analyzed scene, as described by Ullman and Pylyshyn, has a direct analog in human signed languages. Where spoken languages establish the existence of discourse referents with noun phrases, and subsequently use definite pronouns and descriptions to re-identify these referents, signed languages can use a directly visuo-spatial method of keeping track of discourse referents. A user of British Sign Language, for instance, on telling a story involving three participants, will, on introducing them into the discourse, assign them a position in the signing space around him. On referring back to these individuals, he will point to the appropriate spatial position (equivalent to saying "this one" or "that one").

> [In many sign languages] Anaphoric pronouns can only occur following the localization of the referent noun in the location assigned to the pronoun. Nouns articulated in the space in front of the body are, for example, moved to third person space; nouns located on a body part would be followed by an indexing of third person space. This assignment of location to a referent . . . then continues through the discourse until it is changed. To indicate anaphoric reference, the signer indexes the location previously assigned to that referent. . . .
>
> The operation of anaphora . . . can be seen in the following BSL example "The woman keeps hitting the man." In this, the sign MAN is articulated with the left hand, followed by the 'person' classifier, located to fourth person space. The left hand remains in the "person" classifier handshape and fourth person location, while the remainder of the sentence is signed. The sign WOMAN is articulated with the right hand, followed by the "person" classifier, located to third person space. The verb HIT, an agreement verb, is then articulated, moving on a track from the subject (third person) to object (fourth person).[9] (Woll & Kyle 1994, p. 3905)

See also Liddell (1990), McDonald (1994), and Padden (1990). For the sign language recipient, the experience of decoding a signed scene-describing utterance closely parallels the visual act of analyzing the scene itself; in both cases, the objects referred to are assigned to different locations in space, which the recipient/observer marks.

There is a further parallel between linguistic deictic terms and the deictic variables invoked by vision researchers. As we have seen, Pylyshyn postulates "a pool of four or five available indexes," and Ballard et al. (1997) emphasize that most ordinary visually guided tasks can be accomplished with no more than three deictic variables. The deictic terms of natural languages are organized into internally contrastive subsystems: English examples are *here/there, now/then, yesterday/today/tomorrow,* Past-tense/non-Past-tense, *this/that, these/those.* Some languages are slightly richer in their deictic systems than English. Japanese, for instance, distinguishes between three demonstratives, *kono* (close to the speaker), *sono* (close to the listener, or previously referred to), and *ano* (reasonably distant from both speaker and listener); this three-way distinction in demonstrative adjectives is paralleled by three-way distinctions in *kore/sore/are* (demonstrative pronouns) and *koko/*

soko/asoko and *kochira/sochira/achira* (adverbs of place and direction respectively). Spanish likewise makes a three-way distinction in demonstratives, *este/ese/aquel,* with slightly different meanings from the Japanese. There are a few languages with four-way contrasts. Tlingit is one such language. In Tlingit,

> *yáa* "this (one) right here" is clearly "close to *Sp*"; *héi* "this (one) nearby" is characterized by a moderate distance from *Sp* without reference to the *Adr; wée* "that (one) over there" is again not identified by the location of the *Adr;* and *yóo* "that (one) far off (in space or time)," the fourth term, is simply remote from the speech situation. (Anderson & Keenan 1985, p. 286)

Anderson and Keenan mention two other languages, Sre and Quileute, as also having four-way deictic contrasts. They mention one language, CiBemba, with a five-way system, and one, Malagasy, with a seven-way system; frankly, I am skeptical of the claim for seven degrees of contrast along a single dimension in Malagasy. "Systems with more than five terms along the basic deictic dimension are exceedingly rare" (Anderson & Keenan 1985, p. 288).

The extreme rarity of languages providing more than five contrasting deictic terms in any subsystem corresponds nicely to the "pool of four or five available indexes," or visual deictic variables, postulated by Pylyshyn. In an utterance entirely concerning objects in the vicinity of the speech situation, none of which are identified by any predicate/property, there is a limit to how many separate things a speaker or hearer can keep track of, with expressions equivalent to "this one near me," "that one near you," "that one yonder," and so on. Pylyshyn (1989) explicitly relates his FINST devices to the indexical pronouns *here* and *there,* and suggests that FINSTs provide a semantics for such expressions. It is important to note the highly elastic size of the domains appealed to in deixis. Within deictic systems, "near" and "far" are typically relative, not absolute. Hence, within a domain which is all in some sense near the speaker, there nevertheless will still be a distinction between "near" and "far."

The provision by the brain's sensory/perceptual systems of a pool of about four or five variables for ad hoc deictic assignment to objects in the accessible environment, and the separate processes of perceptual categorization of the objects so identified, constitutes an early system for the representation of scenes. This system was based on multiple instances of (or conjunctions of) propositions of the form *PREDICATE(x),* involving up to about four different variables. An example of such a scene-description might be

$$APE(x)\ \&\ STICK(y)\ \&\ MOUND(z)\ \&\ HOLE(w)\ \&\ IN(w,z)\ \&\ PUT(x,y,w)$$

translating to *An ape puts a stick into a hole in a mound.* This translation is given here just for convenience.[10] So far, we have made no move to suggest how such nonlinguistic mental representations came to be externalized in the shared communication system of a community. If we are talking about language at all, it is, so far, only private language. Nevertheless, given the genetic homogeneity of communities of primates, it is highly likely that what happens in the brain of one animal on seeing a scene is represented very similarly in the brains of its fellow troop members. The simply structured internal representations provide a preadaptive platform on which a simple public language could develop.[11]

I have suggested certain parallels between the prelinguistic representation of events (restriction to three to five participants, location of the participants in egocentric space) and features of modern human languages (clause size, limits of deictic systems, anaphora in sign languages). I believe that these features of language can ultimately be traced back to evolutionary precursors in the prelinguistic representations. But it also seems very likely that in the evolution of the language capacity, the human brain has liberated itself from certain of the most concrete associations of the prelinguistic representations. Thus, when a modern human processes a sentence describing some abstract relation, such as *Ambition is more forgivable than greed,* it is unlikely that any specifically egocentric space-processing (parietal) areas are activated. The relation between ancient egocentric visuo-spatial maps and modern features of language is, I would claim, rather like the relationship between ancient thermoregulation panels and wings, a relationship of homology or exaptation. If the ancient structures had never existed, the modern descendants would not have the particular features that they do, but the modern descendants are just that, *descendants,* with the kind of modifications one expects from evolution.

5. Common ground of neuroscience, linguistics, and philosophy

I have made the connection between neural processing of visual scenes and mental representations of propositions as expressed by simple natural language clauses. The same connection is everywhere heavily implicit, though not explicitly defended, in the writing of the vision researchers cited here. In particular, the four terms, "deictic," "indexical," "refer," and "semantic," borrowed from linguistics and the philosophy of language, have slipped with remarkable ease and naturalness into the discussion of visual processing. "Deictic" as a grammatical term has a history going back to the Greek grammarians (who used "deiktikos"; see Lyons 1977, p. 636, for a sketch of this history), indicating a "pointing" relationship between words and things. "Deictic" and "indexical" are equivalent terms. Agre and Chapman (1987) apply the term "indexical" to computational entities invoked by a program designed for fast, efficient, planning-free interaction with a model world. These entities "are not logical categories because they are indexical: their extension depends on the circumstances. In this way, indexical-functional entities are intermediate between logical individuals and categories" (Agre & Chapman 1987, p. 270).[12] The parallels between efficient computing for fast local action and the efficient fast analysis of visual scenes, using deictic or indexical entities, are later taken up by a small but growing number of writers (e.g., Ballard et al. 1995; 1997; Pylyshyn 2000) arguing the advantages of reorientating perceptual and cognitive research along "situated" or "embodied" lines.

Similarly, the term "refer" is typically used in ordinary language, and consistently in the more technical discourse of linguists and philosophers, with a linguistic entity, such as a word, as one of its arguments, and a thing in the world as another argument, as in "*Fido* refers to my dog." Strawson's classic article "On Referring" (Strawson 1950) is all about statements and sentences of ordinary languages; for Searle (1979) and other speech act theorists, referring is a speech act. Linguists prefer to include a third argument, the speaker, as in "He referred to me as *Jimmy.*" Manually pointing to an object, without speaking, might be considered by some linguists and philosophers to be at best a marginal case of referring, especially where the intention is to draw attention of another to the object. But notice how easily this and other originally linguistic terms ("demonstrative," "indexical") are interpreted when applied to a visual, entirely non-linguistic process:

> The visual system . . . needs a special kind of direct reference mechanism to refer to objects without having to encode their properties. . . . This kind of direct reference is provided by what is referred to as a demonstrative, or more generally, an indexical.[13] (Pylyshyn 2000, p. 205)

The central idea involved in linguistic and vision-oriented and activity-oriented uses of the terms "deictic," "indexical" and "refer" is *attention.* In all cases, be it a monkey swivelling its eyes toward a target, an ape grasping for an object, or a human referring to an object with a demonstrative pronoun, the organism is attending to an object. This is the archetypal sense of "refer-"; the linguist's preferred usage of "refer-," involving a speaker, is closer to the archetypal sense than the twentieth century logician's, for whom reference is a relation between words and things, without mediation by any agent's mind. But the linguist's and the philosopher's restriction of "referring" to a necessarily linguistic act misses what I claim is the phylogenetic, prelinguistic origin of referring.

Classically, semantics is said to involve a relation between a representation and the world, without involvement of any user of that representation (e.g., a speaker) (Carnap 1942; Morris 1938; 1946). Thus, the relation of denotation between a proper name and its referent, or between a predicate and a set of objects, is traditionally the concern of semantics. Vision researchers use the term "semantic" with no sense of a relation involving linguistic entities. Jeannerod et al. (1995) identify events in the dorsal stream with pragmatics (though perhaps "praxics" might have been a better term) and events in the ventral stream with semantics:

> In humans, neuropsychological studies of patients with lesions to the parietal lobule confirm that primitive shape characteristics of an object for grasping are analyzed in the parietal lobe, and also demonstrate that this "pragmatic" analysis of objects is separated from the "semantic" analysis performed in the temporal lobe. (Jeannerod et al. 1995, p. 314)

Likewise Milner and Goodale (1995, p. 88) write of the "content or semantics" of nonverbal interactions with the world, such as putting an object in a particular place. Further, "even after objects have been individuated and identified, additional semantic content can be gleaned from knowing something about the relative location of the objects in the visual world" (Milner & Goodale 1995, p. 88). The central idea linking linguists', philosophers', and vision researchers' use of "semantic" is the idea of *information* or *content.* For us modern humans, especially the literate variety, language so dominates our lives that we tend to believe that language has a monopoly of information and content. But of course there is, potentially, information in everything. And since the beginning of the electronic age, we now understand how information can be transmitted, transformed, and stored with wires, waves, and neurons. Information about the relative location of the objects in a visual scene, or about the properties of those objects, represented in a perceiver's brain, has the same essential quality of "aboutness," a relation with an external world, that linguists and philosophers identify with the semantics of sentences. Those philosophers and linguists who have insisted that semantics is a relation between a language and the world, without mediation by a representing mind, have eliminated the essential middleman between language and the world. The vision researchers have got it more right, in speaking of the "semantics" of neural representations, regardless of whether any linguistic utterance is involved. It is on the platform of such neural representations that language can be built.

An evolutionary history of reference can be envisaged, in which reference as a relation between the mind and the world is the original. This history is sketched in Figure 2.

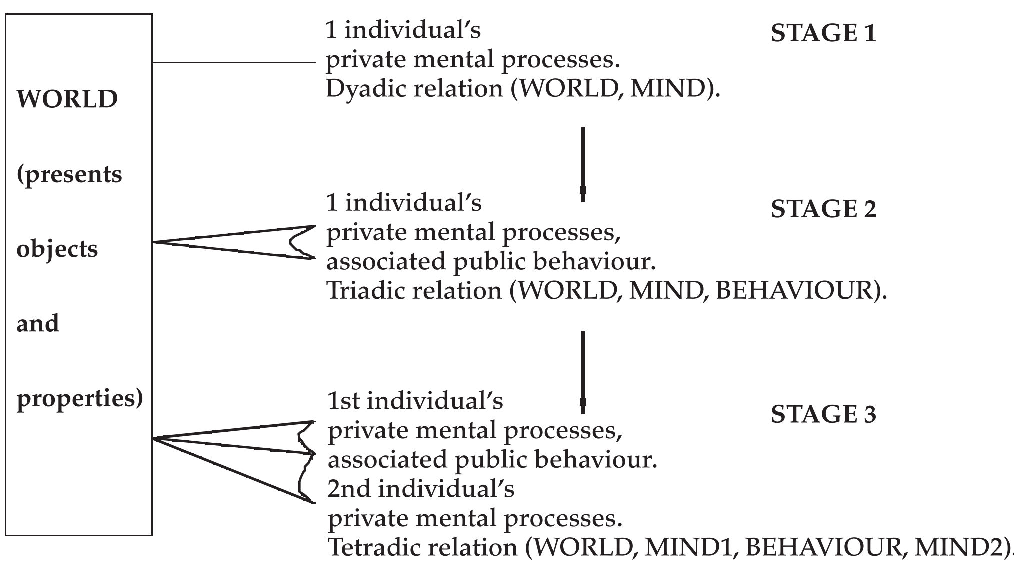

Figure 2. The evolution of reference. The relationship between mental processes and the world is the original and enduring factor. The last stage is successful reference as understood by linguists, and as manifested by people speaking natural languages. The stages may overlap, in that further evolution of one stage may continue to complexify after evolution of a later stage has commenced.

At present, the dual use of such terms as "deictic" and "refer" for both linguistic and visual processes is possibly no more than a metaphor. The mere intuitive plausibility of the parallels between the visual and the linguistic processes is not as good as empirical evidence that the brain in some way treats linguistic deictic variables and visual deictic variables in related ways. Possibly the right kind of evidence could be forthcoming from imaging studies, but the picture is sure to be quite complicated.

6. Wrapping up

6.1. It could have been otherwise

It could conceivably have been otherwise, both from a logical and a biological point of view. Consider, first, alternative biologies. We can conceive of a world in which organisms sense the ambient temperature of their surroundings by a single sensory organ which doesn't distinguish any source of radiant heat. Further, such a creature might have a keen sense of smell, and be able to discriminate between thousands of categorically different smells assailing its smell organ. And the creature might have arrays of light detectors evenly spaced all over its body, all feeding into a single internal organ activated by an unweighted average of the inputs. Such a creature would have no internal representation of objects, but only a set of "zero-place predicates." It could sense "The world outside is in such-and-such a state." Certainly, the higher animals on planet Earth are not like this, but I would be surprised if some lower animals were not somewhat like it. It just happens to be the case that the laws of physics, chemistry, and biology conspire to produce a world containing discrete categorizable objects, and so, not surprisingly, but *not logically necessarily,* advanced creatures have evolved ways of dealing with them.

An alternative logic is also easily conceivable, in which there is no predicate-argument structure. It already exists in the form of the propositional calculus, typically introduced in logic textbooks as a simple step towards the more "advanced" predicate calculus. A propositional calculus, with no predicate-argument structure, would be all that is needed by the creature described in the previous paragraph.

Here is a final thought experiment. A "Turing robot" is entirely conceivable as a working automaton, capable of navigating and surviving in a complex world. Instead of reading a character on a tape, the Turing robot "reads" a patch of the world in front of it, matching the input to some monadic symbol occurring in the quadruples of its instruction set. Instead of shifting the tape to right or left, it shifts itself to an adjacent patch of world, and it can act, one monadic action at a time, on the patch of world it is looking at. Given a complex enough instruction set, such a robot could replicate any of the complex computations carried out by an advanced real live creature successfully negotiating the world. The Turing robot's hardware, and the individual elements of its software instruction set, the basic quadruples, contain nothing corresponding to predicate-argument structure, though it is probable that we could interpret some higher-level pattern or subroutine in the whole instruction set as somehow corresponding to predicate-argument structure. The dorsal/ventral separation in higher mammals is, I argue, an evolved hardware implementation of predicate-argument structure.

6.2. Falsifiability

This article is an instance of reductionism. It takes two previously unrelated fields, logic and neuroscience, and argues that what logicians are really dealing with, whenever they appeal to predicate-argument structure, has a basis in neural processing. This in no way minimizes the validity of studies in logic; rather it enhances their validity. Biologists working with Mendelian genes without knowledge of DNA were doing valid work. "Abstract" work on the structure of human thought, and its relationship to language, must continue. But as long as we recognize that the object of study, both in logic and in linguistics, has a psychological basis, one of us should also work on bridging the gap between theoretical studies couched in logico/linguistic terminology and empirical studies in psychology and neuroscience. Only those who view logical and linguistic structure as Platonic, in some way existing independently of human minds, can ignore psychology and neuroscience.

Can a reductionist argument be falsified? Yes. Some proposed reductions are just plain wrong, some are well justified, and some are partly right. What justifies a reductionist argument is the goodness of fit between the two independently established theories. The present argument would be invalidated if it could be shown that any of the following apply:

The canonical arguments of predicates in logic do not denote individual objects.

Canonical predicates in logic do not denote properties.

The dorsal stream processes properties at least as much as the ventral stream.

The ventral stream plays a large role in drawing attention to objects.

I concede that an extreme version of my reductionist proposal *is* falsified in many ways, because, on the logical side, for example, formal semanticists often use nonobject-denoting terms as arguments of predicates, and on the neurological side, for example, some detection of properties is achieved by the dorsal stream. So the fit between the practices of logicians and formal semanticists with predicate-argument structure and the neural facts is not quite perfect. But, I claim, there is enough of a clear parallelism between the two domains to indicate that neuroscience has revealed facts which significantly inform the domain that logicians and formal semanticists traditionally deal with. Here again I mention that the brain is vastly more complex than even the most baroque of logical formalisms, and that one should expect complexities arising from brain studies that logical studies simply do not relate to. A logical formalism relates to the brain in the same way as a road map relates to a real place.

6.3. Then, now, and next

The neural correlates of *PREDICATE(x)* can be found not only in humans but also in primates and probably many other higher mammals. Thus, as far as human evolution is concerned, this form of mental representation is quite "primitive," an early development not unique to our species. It can be seen as building on an earlier stage (evident, for example, in frogs) in which the only response to an attention-drawing stimulus was some immediate action. A fundamental development in higher mammals was to augment, and eventually to supplant, the immediate motor

http://dx.doi.org/10.1017/S0140525X03380070

responses of a sensorimotor system with internalized, judgmental responses which could be a basis for complex inferential processes working on material stored in long term memory. Rather than "If it moves, grab it," we begin to have "If it catches your attention, inspect it carefully and figure out what do to with it," and later still "If you notice it, remember what is important about it for later use."

Simple early communicative utterances could be reports of a *PREDICATE(x)* experience. For example, the vervet chutter could signify that the animal is having a *SNAKE(x)* experience, that is, has had its attention drawn to an object which it recognizes as a snake. Primitive internal representations, I have claimed, contain two elements, a deictic variable and a categorizing predicate. Nowhere in natural nonhuman communication do we find any two-term signals in which one term conveys the deictic element and the other conveys the mental predicate. But some simple sentences in some human languages have just these elements and no other. Russian and Arabic provide clear examples.

eto	čelovek	
DEICTIC	MAN	"This is a man." (Russian)
di	sahl	
DEICTIC	EASY	"That is easy." (Egyptian Arabic)

Even if the internal representations of animals are structured in the *PREDICATE(x)* form, there would be no evolutionary pressure to structure the corresponding signals into two parts until the number of possible mental combinations of predicates and variables exceeded the total number of predicates and variables, counted separately (Nowak et al. 2000). If the category of things that are pointed to in a given direction is always the same, there is no pressure for the signal to differentiate the direction from the category.

I have argued that *PREDICATE(x)* is a reasonable schematic way of representing what happens in an act of perception. It is another step, not taken here, to show that a similar kind of logical form is also appropriate for representing stored episodic memories. A form in which only individual variables can be the arguments of predicates might be too restrictive. Here, let me, finally, mention the "Aristotle problem." Aristotle and his followers for the next two millennia took the basic semantic representation to be Subject + Predicate, where the same kind of term could fill both the Subject slot and the Predicate slot. Thus, for example, a term such as *man* could be the subject of *The man died* and the predicate of *Plato is a man.* Kant's characterization of analytic judgements relies on subject terms being of the same type as predicate terms. "Analytical judgments express nothing in the predicate but what has been already actually thought in the concept of the subject, though not so distinctly or with the same (full) consciousness". (Kant 1905, translation of Kant 1783).[14] FOPL is more distanced from the surface forms of natural languages, and the same terms cannot be both arguments (e.g., subjects) and predicates. It remains to provide an explanation for the typical structure of modern languages, organized around the Noun/Verb dichotomy. I suspect that an explanation can be provided in terms of a distinction between predicates which denote invariant properties of objects, such as being a dog, and more ephemeral properties, such as barking. But that is another story.

NOTES

1. The logical formula is simplified for convenience here.

2. A complication to this picture arises from work on the recognition of facial expressions by blindsight patients (de Gelder et al. 1999; 2000; Heywood & Kentridge 2000; Morris et al. 1999). Facial expressions are complex and are generally thought to require considerable higher-level analysis. Yet detection of facial expressions (e.g., sad, happy, fearful, angry) is possible in some blindsight patients, suggesting that some aspects of this task also are performed via a pathway that, like at least one dorsal pathway, bypasses primary visual cortex.

3. Belin and Zatorre (2000) suggest that the dorsal auditory pathway is involved in extracting the verbal message contained in a spoken sentence. This seems highly unlikely, as parsing a sentence appeals to higher-level lexical and grammatical information. The evidence they cite would only be relevant to the early pressure-sequence-to-spectrogram stages of spoken sentence processing.

4. Landau and Jackendoff (1993) is a more detailed version of Jackendoff and Landau (1992); I will refer here to the later paper, Landau and Jackendoff (1993).

5. Bertrand Russell at times espoused the view that particulars are in reality nothing but bundles of properties (Russell 1940; 1948; 1959). See also Armstrong (1978). There is also a phenomenalist view that "so-called material things, physical objects, are nothing but congeries of sensations" (Copi 1958).

6. Egly et al. (1994) state:

> We found evidence for both space-based and object-based components to covert visual orienting in normal observers. Invalid cues produced a cost when attention had to be shifted from the cue to another location within the same object, demonstrating a space-based component to attention. However, the costs of invalid cues were significantly larger when attention had to be shifted an equivalent distance and direction to part of another object, demonstrating an object-based component as well. (p. 173)

This again conflates attention-shifting, a preattentive (and post-attentive) process, with attention itself. These experiments relate only to attention-shifting, as the title of Egly et al.'s (1994) article implies. (Further, it would be interesting to know whether the distribution of RTs for the invalidly cued "within-object" attention shifts was in fact bimodal. If so, this could suggest that subjects were sometimes interpreting the end of a rectangle as a *different object* from the rectangle itself, and sometimes not. In this case, the responses taken to indicate a space-based process could in fact have been object-based.)

7. There is presumably a complex ecological balance between the information carried by a mental predicate and its frequency of use in the mental life of the creature concerned. Complex relations, if occurring frequently enough, might be somehow compressed into unitary mental predicates. An analogous case in language would be the common compressing of *CAUSE(a, PRED(b))* into a form with a single causative verb.

8. The claim in the text is not about memory limitations involved in parsing linguistic strings; it is about how many arguments the elementary propositions in the mind of a prelinguistic creature could have.

9. Both third and fourth person space in BSL are like available pronouns for entities being signed about, other than the speaker or hearer. It is not that BSL has four grammatical persons in the sense that English has three (1st – speaker, 2nd – hearer, 3rd – all other entities).

10. This formula, like any FOPL formula, conveys no temporal transitions. Tense logic is more complicated than FOPL.

11. See Batali (2002) for a computer simulation of the emergence of public language from representations of exactly this form.

12. Agre and Chapman (1987) do not, as stated by Ballard et al. (1995), use the term "deictic."

13. Indeed, this quoted sentence contains the stem "refer-" four times, three times alluding to a visual process and once to a linguistic convention; probably few readers remark on the coincidence as in any way disturbing.

14. Notice in this quotation from Kant a faint forerunner of the idea developed in this article, that predicates are associated with processes more accessible to consciousness than arguments.

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

Probability rather than logic as the basis of perception

Thomas J. Anastasio

Department of Molecular and Integrative Physiology, Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801. **tja@uiuc.edu http://csn.beckman.uiuc.edu/**

Abstract: Formal logic may be an inappropriate framework for understanding perception. The responses of neurons at various levels of the sensory hierarchy may be better described in terms of probability than logic. Analysis and modeling of the multisensory responses of neurons in the midbrain provide a case study.

I find merit in Hurford's basic idea that the ventral stream, or "what" cortical pathway, somehow classifies the objects or events to which attention is directed by the dorsal, "where" pathway. What I question is the form Hurford proposes for the nature of this classification. Hurford suggests a process analogous to the evaluation of a statement in formal logic. I suggest the process could be better understood as the evaluation of a probability.

That brain processes can be modeled using logical formalisms is an enduring theme. The desire to understand thought processes motivated the logical methods developed by George Boole. Famous work by McCulloch and Pitts (1943) demonstrated that neural networks could carry out logical operations. More recent work shows how cognitive processes could be built up from networks of neural elements that can learn basic logical functions (Valiant 1994). Whereas the brain could implement logical operations in principle, convincing evidence that it does so in practice is lacking. For example, the initial hope that brain-like intelligence could be created using artificial systems based on logic (e.g., McCarthy 1968; Newel 1982) has been lost.

Other concepts have had more success in providing insight into brain function. Barlow (1969; 1972) suggested that neurons throughout the sensory hierarchy could be thought of as feature detectors, the responses of which are proportional to the probability that their trigger feature is present. Barlow's original model was cast in terms of classical statistical inference. More recent incarnations of this idea involve Bayesian methods. Probabilistic models are distinctly better then those based on formal logic for understanding the response properties of sensory neurons. Research on multisensory neurons in the midbrain provides a case in point.

Multisensory interactions were first described by Newman and Hartline (1981), for neurons in the rattlesnake optic tectum that combine input from the visual and infrared pit-organ systems. Newman and Hartline categorized the responses of multisensory tectal neurons to visual and infrared stimuli presented separately (modality-specific) or together (cross-modal), near receptive field centers in both modalities. The ideal OR neuron had both modality-specific and cross-modal responses, whereas the ideal AND neuron had only cross-modal responses. They seemed to compute Boolean logical functions on their inputs. The analogy between the responses of multisensory tectal neurons and Boolean logical operators, however, could only be taken so far.

The cross-modal response of OR neurons could be larger than either of the modality-specific responses, and even larger than their sum. The modality-specific responses of AND neurons could be nonzero. Other neurons could not be fit into a Boolean scheme at all. For example, the responses of ENHANCED tectal neurons to a stimulus of one modality could be increased by a stimulus of another modality that was ineffective by itself. The responses of all types were significantly magnitude dependent. It would not be possible, on the basis of the data on multisensory neurons in the rattlesnake tectum, to develop a satisfying description of their response properties in terms of Boolean logic.

Later work by Meredith and Stein (1983; 1986; for review, see Stein & Meredith 1993) provided a more general view of multisensory responses. They studied multisensory neurons in the deep layers of the mammalian superior colliculus (DSC), and described ENHANCEMENT as *any* augmentation of the response to stimulation of one sensory modality by the presentation of a stimulus of another modality. These responses are also magnitude dependent. Percent enhancement is larger when modality-specific responses are smaller. That property, known as INVERSE EFFECTIVENESS, provides the key to a model that can unify findings on multisensory interactions. This model is based not on logic, but on probability.

We modeled INVERSE EFFECTIVENESS on the hypothesis that multisensory DSC neurons use their inputs to compute the probability that a target, defined as a stimulus source, has appeared in their receptive fields (Anastasio et al. 2000). Specifically, we propose that DSC neurons compute $P(T = 1|S)$, where P is probability, S is sensory input of one or more modalities, and T is the target ($T = 1$, target present; $T = 0$, target absent). This conditional probability can be computed as a posterior probability using Bayes' rule: $P(T = 1|S) = P(S|\mathrm{T} = 1)P(T)/P(S)$. By equating posterior probabilities with the responses of multisensory DSC neurons, the Bayes' rule model can simulate INVERSE EFFECTIVENESS. If a modality-specific stimulus is large, it provides overwhelming evidence of a target. The posterior probability of a target would be close to 1, and a stimulus of another modality would not increase it much. However, if a modality-specific stimulus is small, the posterior probability of a target would be close to 0. Integrating a stimulus of another modality would dramatically increase the probability that a target has appeared. The correspondence between posterior probabilities and the multisensory responses of DSC neurons strongly supports the hypothesis that DSC neurons compute the probability of a target given their multisensory inputs.

Of course, the brain does not compute by using probability distributions, but through synaptic weights and neural activation functions. Borrowing techniques from the field of statistical pattern classification (Duda et al. 2001), we developed simple neural models that are capable of computing posterior probabilities exactly, given well-described input distributions (Patton & Anastasio 2003). These models simulate ENHANCEMENT and INVERSE EFFECTIVENESS. For input distributions that are not well described, posterior probabilities can be estimated using neural networks (Bishop 1995). Thus, artificial neurons and networks are well suited to the computation of posterior probabilities, and it is reasonable to suppose that the brain is also.

Multisensory neurons in cortex have response properties similar to those in the DSC (Stein & Wallace 1996), and the hypothesis that multisensory neurons compute posterior probabilities could be extended to other brain regions. Cortical neurons, in general, could be thought of as feature detectors that compute the posterior probability that their feature has been detected, given their inputs. In his target article, Hurford suggests that neurons in the ventral ("what") stream of cortical processing evaluate the logical statement PREDICATE(X). X stands for an object or event in the environment that elicits sensory input *S*, the location of which is marked for attention by the dorsal ("where") stream of cortical processing. Presumably, neurons in the ventral stream would have two states, active or silent, corresponding to the "true" or "false" values of PREDICATE(X). By way of a concrete example, we might consider a hypothetical neuron in the ventral stream that evaluates APPLE(X), which would fire neural impulses at some fixed rate if the currently attended sensory input *S* corresponds to an apple, and would be silent otherwise. One problem with this scheme is that cortical neurons are not two-state elements but show graded responses to their inputs. Another, more serious, problem is that sensory inputs, being neural, are stochastic and therefore uncertain to some extent. It is hard to see how neurons in the ventral stream, or anywhere else in the brain, could ever be completely certain of exactly what has elicited their current pattern of input. Perhaps a better way to model the responses of such a neuron is $P(X = \text{APPLE}|S)$. The activity of the neuron would then vary from zero up to some maximal level of firing, which would be proportional to the probability that object *X*, eliciting sensory input *S*, is an apple. The computation of probabilities seems a more realistic basis for perception than the evaluation of statements in logic.

Prelinguistic agents will form only egocentric representations

Michael L. Anderson[a] and Tim Oates[b]

[a]*Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742;* [b]*Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21250.*
anderson@cs.umd.edu oates@cs.umbc.edu
http://www.cs.umd.edu/~anderson
http://www.csee.umbc.edu/~oates/

Abstract: The representations formed by the ventral and dorsal streams of a prelinguistic agent will tend to be too qualitatively similar to support the distinct roles required by *PREDICATE(x)* structure. We suggest that the attachment of qualities to objects is not a product of the combination of these separate processing streams, but is instead a part of the processing required in each. In addition, we suggest that the formation of objective predicates is inextricably bound up with the emergence of language itself, and so cannot be cleanly identified with any prelinguistic cognitive capacities.

In his search for the neural basis of the simple logical structure *PREDICATE(x),* Hurford focuses on the basic cognitive capacities of advanced, but prelinguistic agents – for he follows Batali (2002) in believing that language can be developed out of such simple logical structures. Not having any such agents to study directly, Hurford instead considers the capacities of higher primates, abstracting away from their linguistic abilities to uncover two basic perceptual processing systems – the dorsal and ventral streams (henceforth DS and VS) – which he suggests provide the basic components of *PREDICATE(x)* structure. However, he does not carry this thought experiment through consistently; and, failing to consider matters from the perspective of such a prelinguistic agent, he is led to interpret the information delivered by these perceptual processing streams in postlinguistic terms, thereby in a subtle way assuming what he is trying to prove. When things are instead considered from the standpoint of such an agent, it no longer looks as if the DS and VS provide attractive candidates for the separate components of *PREDICATE(x)* structure.

Hurford rightly emphasizes that the deliverances of the DS – the "where" pathway that provides information about the location and size and shape of an object – are cast in "egocentric" terms. The DS is a specialized perceptual processing system that represents information in a form optimized for calculating and directing motor responses aimed at an object in virtue of its location, orientation, and spatial extent. This information is used to guide such things as the orientation of sense organs for optimal perception, perceptual tracking, reaching, and grasping. Thus, the natural way to characterize what the agent knows in virtue of DS representations (what the information means to the agent) is in terms of egocentric spatial coordinates: Where it is in relation to the agent, and what might be done to get the agent-object relation into a preferred state. One might say that the DS places objects in an egocentric visuomotor space, or an egocentric action field, and the object is thereby presented to the agent in these terms.

This characterization of the function of the DS is largely in accord with Hurford's – but what drives Hurford's account is the supposed *contrast* between the egocentric "motor-oriented" information given by the DS, with the "cognitive" – and therefore in some sense more objective – information said to be delivered by the VS. But in the individual, prelinguistic, and thus (one might say) *functionally solipsistic* agents Hurford describes, this contrast is untenable. Like the DS, the VS is a specialized perceptual processing system, but in this case it is optimized for representing information about the *look* rather than the *location* of an object. Just as with the DS, the VS representations are used to select and direct appropriate motor responses; and just as with the DS, the most natural way to characterize what the agent knows in representing this information is what the object means to it – to its utility, goals, survival – and what it might appropriately do in response. VS representations likewise place the object in a visuomotor action field, useful for calculating responses based on the differences between individual objects as opposed to individual locations. To imagine instead that encountered objects are represented in terms of objective features or abstract qualities is to import into the VS the kind of representational scheme appropriate for language and logic, but of no use to agents whose primary concerns are individual and behavioral but not communal or communicative.

So what we, in fact, have in the case of the individual agents Hurford describes are two specialized processing mechanisms that, although optimized for representing different aspects of perceptual information, are both nevertheless engaged in interpreting that information in conceptual terms suited to the selection and direction of appropriate motor responses. It does not seem that the products of these two visuomotor control systems lend themselves to natural combination in the form Hurford needs. Indeed, there seem to be substantial gaps between what these pathways deliver – the egocentrically presented features of objects – and *PREDICATE(x)* structure. We next describe two of these gaps and identify possible approaches to bridging them.

Consider first the fact that the representations formed by the VS and DS are *egocentric.* Language is useful only to the extent that it enables agents to share meaning with one another, but for meaning to be shared, it must be objective, not subjective or idiosyncratic. An isolated agent capable of forming prelinguistic concepts, or *predicates,* from the information delivered by its VS will form just those concepts that help this one agent survive in whatever environment it finds itself. These concepts will be cast within a single, agent-centered frame of reference, meaningful only from its own individual perspective. But suppose this agent discovers others of its kind. Through repeated interactions with these new agents, objective features of the world – those features commonly available and salient to others – can be identified and thus used to form the concepts (predicates) that serve as the semantic basis of language. The computational model of language evolution developed by Luc Steels (Steels 1997) leverages this idea of repeated interactions to separate the objective from the subjective and thereby evolve a stable, shared lexicon. Genuinely objective predicates, and the representational schemes that support them, arise only as the result of the formation of such shared, stable, intersubjective representation systems.

Second, it's clear that in *PREDICATE(x)* the two components – *PREDICATE()* and *x* – are qualitatively different. But if both processing streams are delivering the egocentrically presented features of objects, then neither stream seems properly specialized for producing *objects* over *properties.* Rather, if *both* pathways can be said to return information about objects – such as the distance to an object in the case of the dorsal pathway or the color of an object in the case of the ventral pathway – then it seems that *both* pathways must have access to (or contain) a neural mechanism that individuates objects in the visual field, making it possible to bind deictic markers and extract information about, for example, distances or colors. Given this mechanism, 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).* Indeed, it seems that the data presented in (Goodale et al. 1994) support this view, for the patient R.V. was able to discriminate objects despite apparent damage to DS processing.

Predicates: External description or neural reality?

Michael A. Arbib

Computer Science Department, Neuroscience Program and USC Brain Project, University of Southern California, Los Angeles, CA 90089-2520.
arbib@pollux.usc.edu http://www-hbp.usc.edu/

Abstract: Hurford argues that propositions of the form PREDICATE(x) represent conceptual structures that predate language and that can be explicated in terms of neural structure. I disagree, arguing that such predicates are *descriptions* of limited aspects of brain function, not available as representations in the brain to be exploited in the frog or monkey brain and turned into language in the human.

Note: The numbered paragraphs relate to the corresponding sections of the target article; unnumbered paragraphs contain my comments.

(Section 1.2.) The basic ontological elements are whole events or situations and the participants of these events. The event described by *A man bites a dog* could be represented as

$$\exists e, x, y, bite(e), man(x), dog(y), agent(x), patient(y) \qquad (1)$$

I don't think this works. We need to replace *agent x* by *agent(x, e)* to indicate in which event *x* plays the stipulated role; similarly for *y.* For Hurford, the discussion of episodes is an aside to his concentration on 1-place predicates, but I suggest that the crux for a prelinguistic representation is the event and the "action-object frame" *A(x,y)* – agent *x* is doing *A* to object *y* – and its variations. Rizzolatti and Arbib (1998) examined "whether a 'prelinguistic grammar' can be assigned to the control and observation of actions. If this is so, the notion that evolution could yield a language system 'atop' the action system becomes much more plausible" (p. 191).

This talk of a prelinguistic grammar was not meant to imply that gestures may be a primitive form of grammar, for our approach was semantic rather than syntactic:

> We might say that the firing of "mirror" F5 neurons is part of the code for a declarative case structure, for example,
>
> Declaration: grasp-A(Luigi, raisin)
>
> which is a special case of grasp-A(agent, object), where grasp-A is a specific kind of grasp, applied to the raisin (the object) by Luigi (the agent . . . *this is an "action description," not a linguistic representation* (Rizzolatti & Arbib 1998, p. 192. Emphasis added).

Being able to grasp a raisin is different from being able to say, "I am grasping a raisin," and the neural mechanisms that underlie the doing and the saying are different. However, the case structure lets us see a commonality in the underlying representations, thus helping us understand how a mirror system for grasping might provide an evolutionary core for the development of brain mechanisms that support language.

(Section 2.) Representations of the form PREDICATE(x) are taken to stand for the mental events involved when a human attends to an object in the world and classifies it perceptually as satisfying the predicate in question.

More specifically, the notion is that a person may attend to a limited number of objects, and x then stands as an index for one of those objects. Thus, a scene might be represented by a conjunction

$$P1(X1)\ \&\ P2(X2)\ \&\ P3(X3)\ \&\ P4(X4) \qquad (2)$$

where each Xj indexes some region of the scene, and $Pj(Xj)$ indicates that the object at that location possesses property Pj. This leads to another point which (I think) weakens Hurford's critique of Rizzolatti and Arbib (1998):

(Section 4.) An example of a scene-description might be

$$APE(x)\ \&\ STICK(y)\ \&\ MOUND(z)\ \&\ HOLE(w)\ \&\ IN(w,z)\ \&\ PUT(x,y,w) \qquad (3)$$

translating to *An ape puts a stick into a hole in a mound.*

The inclusion of PUT(x,y,w) in (3) reinforces the point that Hurford's focus on unary predicates does not do justice to describing animals which perceive to act, with acts dependent on relations between objects. The key question remains: "How do we go from predicates that we may use to describe internal behavior to neural representations that themselves abstract from the activity levels and parameterizations of schemas and their underlying neural networks, and instead provide abstractions that may in turn be refined to yield the cognitive and semantic forms that drive the production and perception of the phonological forms of language?"

In discussing the possible neural basis of (2), Hurford (sect. 4) cites papers from 1984 onward. However, I would claim some priority in this area with *the slide-box metaphor* (Arbib 1972; Didday & Arbib 1971): In the days before computer graphics, movie cartoons were drawn using *cels,* which I there called *slides.* Because the cartoon might run for seconds without the background changing, one may draw this background just once. In the middle ground, there might be a tree about which nothing changes for a while except its position relative to the background. It could thus be drawn on a separate slide and repositioned as needed. In the foreground, key details might change for each frame. The slides could then be photographed appropriately positioned in a slide-box for each frame, with only a few parameter changes (including minimal redrawing) required between successive frames. The slide-box metaphor suggested that a similar strategy might be used in the brain, with long-term memory (LTM) corresponding to a "slide file" and working or short-term memory (STM) corresponding to the "slide-box." The act of perception was compared to using sensory information to update slides already in the slide-box and to retrieve other slides as appropriate, experimenting to determine whether a newly retrieved slide fits sensory input "better" than one currently in the slide-box, which, in the brain, corresponds to a mass of neural tissue linking sensory and motor systems. A crucial point was that retrieval of a slide provided access to a wealth of information about the object it represented, including appropriate courses of action.

I cite this background to stress that (3) is a pale approximation of the slide-box metaphor, which is in turn a pale approximation to the multilevel modeling methodology that unifies the functional schemas of schema theory (Arbib 1981; Arbib et al. 1998) with the dynamics of detailed neural networks. For example, one *schema* in the visuomotor system of a frog (Arbib 1987a) might correspond to a pattern of neural activity signaling the likelihood of a small moving object in a region x_1 of the visual field, another schema might signal the likelihood that a large moving object is moving with velocity v in region x_2, while a third might indicate the likelihood that a barrier of extent w is located around region x_3. Thus, rather than being predicates that return 0 or 1, they are functions or likelihood distributions over a multidimensional parameter space. Moreover, the frog's actual course of action (the "choice" of motor schema to guide action, and the setting of control parameters for that action) cannot be directly inferred from these schemas, but depends rather on the interplay of the activity of their neural instantiations as they play upon the brainstem, determining whether the frog will snap at its apparent predator or jump to escape an apparent predator (modulating its direction of escape on the perceived trajectory of the predator), and whether or not it will attempt to detour around a barrier in doing so.

In summary, (2) is a fine answer to the question, "What objects does the animal see, and where does it see them?" and Hurford provides an interesting analysis of relevant neural data. Moreover, I think it is useful to debate whether representations in the ventral stream are "more prelinguistic" than those in the dorsal stream. But I answer the question of my title – "Predicates: External description or neural reality?" – by saying that the predicates like (2) are, in general, our external descriptions, not the animal's neural reality. It is a highly evolved skill of humans to be able to name an indicated object, and I suggest that PREDICATE(x) is best seen as a description of human naming behavior, rather than as a conceptual structure that is part of the causality of neural circuits preexisting language.

ACKNOWLEDGMENT
Preparation of this commentary was supported in part by a Fellowship from the Center for Interdisciplinary Research of the University of Southern California.

Afferent isn't efferent, and language isn't logic, either

Derek Bickerton

Department of Linguistics, University of Hawaii, Honolulu, HI 96822.
bickertond@prodigy.net

Abstract: Hurford's argument suffers from two major weaknesses. First, his account of neural mechanisms suggests no place in the brain where the two halves of a predicate-argument structure could come together. Second, his assumption that language and cognition must be based on logic is neither necessary nor particularly plausible, and leads him to some unlikely conclusions.

Hurford is to be commended for attempting to root his analysis of linguistic primitives in the workings of the brain. But there are some serious problems with his approach.

He demonstrates interesting parallels between, on the one hand, the dorsal and ventral pathways that carry information about the "where" and the "what" of objects, and on the other hand, the two halves, predicate and variable, of simple propositions. But a parallelism, intriguing though it may be, falls short of an explanation or even a description.

The dorsal and ventral streams carry information about real-world objects along afferent fibers that lead to very different areas of the brain. A proposition of the form PREDICATE(x) consists not of raw information about external objects but of internal representations of objects and their properties. This already suggests that a lot is missing from Hurford's account. His abstract tells us that "PREDICATE(x) is a schematic representation of the brain's integration of the two processes of delivery by the senses," but of course what the brain integrates is not the raw sensory feels delivered to it but the already stored abstract representations that it decides (correctly or otherwise) are a match for those feels.

Moreover, for these representations to be translated into anything like a predicate–argument structure, 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.

Hurford might want to argue that such considerations lie beyond his scope, that he merely wished to demonstrate the existence of a distinction necessary, though not in itself sufficient, for even the most basic processes of language or thought. However, one can go still further back into his argument and question whether PREDICATE(x) is as fundamental to either thought or language as he supposes.

Hurford's insistence on approaching language from a logical point of view begins to look questionable when he doubts whether proper names existed in earlier forms of language. This doubt is based on the belief that "control of a proper name in the logical sense requires Godlike omniscience." So it may be "in the logical sense," but what were our remote ancestors most concerned about, getting their FOPL straight or telling one another interesting things? I doubt (contra Dunbar 1993) that language was born for gossip, but gossip was surely a major function of language, or even protolanguage, from quite early on. How can you gossip without names for the people you're talking about?

Hurford's supporting arguments are quite weak. It is immaterial whether animals with discriminatory skills less subtle than ours can be fooled into thinking a cardboard cutout is their mother: They have a clear concept of a specific individual, "Mother" – but it is concepts, not things in the world, that matter here – and if they wrongly identify that individual once in a while, so what? Situations where A has sex with B under the impression that he or she is having sex with C have been a staple of farce for millennia; so, logically speaking, Hurford should deny us too the right to have proper names. Tribes claimed still not to use proper names do of course use them. What determines whether something is a proper name is not its internal structure but how it is used. An expression like "knocked the hut over" is (part of) a sentence in "Last night the wind [knocked the hut over]," but in the context of "[Knocked the hut over] seduced your wife last night," it's every bit as much a proper name as is "Jim Hurford."

Further doubts about the logic–language connection spring from the division, in Figure 1, of nouns into the classes "Proper" (said to be arguments) and "Common" (said to be predicates), and are reinforced by the statement in section 2.3.1 that "the vast majority of words in a language correspond to predicates." The vast majority of words are common nouns, which in the vast majority of sentences are arguments, not predicates; they occur as predicates only in sentences of a type seldom uttered by nonlogicians ("Socrates is a man") that were probably rarer still in the dawn of language.

To assume, as Hurford seems to, that such predications form invisible but ineradicable subparts of normal sentences takes us back to the heyday of generative semantics, when to derive a simple sentence like "Floyd broke the glass" required the integration of a dozen or so clauses, including things like "There is someone," "Someone is Floyd," "There is something," and "Something is a glass." But the fact that you can transform the simplest sentence into such components if you try hard enough has no necessary connection with how those sentences actually get constructed (as most syntacticians quickly realized).

A more plausible (and more parsimonious) scenario might go something like this. 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. If it is, the existence of something out there matching something in here is simply presupposed. (In logic you may have to assert your presuppositions, but that's no reason to assume that the brain has to do it that way). Thoughts, or sentences, can then be assembled using a handful of predicates – verbs, prepositions, and the like, that take common or proper nouns indiscriminately as their arguments.

Whether some such scenario, or Hurford's, is nearer the truth

http://dx.doi.org/10.1017/S0140525X03380070

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 subserve 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 rectangular 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**
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 attention. 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**
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 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 occasional encounters at conferences, and I meet at one conference a person who looks very much like Jim (similar height, hair colour, voice quality, and so on). However, this person is Tim, who has agreed to impersonate Jim in order to bamboozle unwitting colleagues. I may well be taken in for a few minutes, or indeed for the whole duration of the conference. Does this mean that I have no concept of Jim Hurford as an individual with spatio-temporal continuity and a unique life-history – a proper-nameable individual, in other words? Clearly it does not. Likewise, the terns' befuddlement by the loudspeaker does not show that they lack any concept of at least some other terns as individuals, and hence it does not show that they lack the mental underpinning to use proper names, supposing that they were linguistically equipped to do so.

In support of his view that proper names are linguistically "late," Hurford points out that some proper names are derived from definite descriptions, such as *Baker, Wheeler, Newcastle.* But one might just as well argue that the fact that signing chimps can use ASL signs that are classified as proper names ("ROGER," "WASHOE," etc.), even if such signs are not derived from definite descriptions, establishes that our protolinguistic ancestors – indeed, the common ancestors of chimps and humans – must have been mentally equipped to use proper names too. So it is risky to derive conclusions about proto-cognition from proper name etymologies and usage.

Hurford mentions the language *Matsigenka.* This is spoken in a community in which there are genuinely no personal names, the few individuals that a person interacts with regularly being identified solely by kinship terms: "father," "patrilineal same-sex cousin," and so on. But this surely demonstrate the very opposite of what Hurford thinks. I would be surprised if a Matsigenka speaker treats as a single individual all his or her relatives to which the same kinship term applies. That is, I assume that, if a Matsigenka speaker has two relatives, both of whom fall under the term glossed "patrilineal same-sex cousin" (for example), he or she will nevertheless be aware that they are different individuals and will treat them as such. If I am wrong in this assumption, then Hurford's case is supported. However, if I am right, what it shows is that the unavailability of proper names has no bearing on the ability to recognise entities to which proper names could appropriately be applied, if the necessary framework (social as well as linguistic) permitted that.

The reason why Hurford is so keen to establish this point, it seems, is an assumption that there must be some stage of linguistic evolution at which (proto)syntax behaved in a fashion that reflected more closely than it does now the way in which predicate-argument structure works in logic. This assumption, if correct, opens up the possibility that, as Hurford puts it, "the dorsal/ventral separation in higher mammals is . . . an evolved hardware implementation of predicate-argument structure." However, I do not share his assumption. Even if it is true that proper names are complex to handle in first-order predicate logic, there is no reason to suppose that what came early in the mental representations of (proto)humans is the same as what is basic in logical terms. Whatever one thinks of evolutionary psychology, it has drawn attention to the fact that mental tasks that are simple, in some logical sense, are not necessarily done well by most humans, whereas there are "harder" tasks that we handle with ease. That is likely to be true of how language arose, too. I have expressed elsewhere scepticism about whether the neurophysiology of vision can explain much in language (Carstairs-McCarthy 1999, pp. 90–91). Hurford's exploration is interesting, and he is evidently well-informed on brain neurophysiology, but he has not made me any less sceptical.

Hurford's partial vindication of classical empiricism

Fiona Cowie

Division of Humanities and Social Sciences, California Institute of Technology, Pasadena CA 91125. **cowie@hss.caltech.edu**

Abstract: Hurford's discussion also vindicates the classical empiricist program in semantics. The idea that *PREDICATE*(x) is the logical form of the sensory representations encoded via the dorsal and ventral streams validates empiricists' insistence on the psychological primacy of sense data, which have the same form. In addition to knowing the logical form of our primitive representations, however, we need accounts of (1) their contents and (2) how more complex thoughts are derived from them. Ideally, our semantic vocabulary would both reflect the psychological "primitiveness" of these representations and make clear how more complex representations derive from them.

Hurford presents himself as vindicating the semantic assumptions of logicians (as to the primacy of predication) and offering insights into the evolution of language (and particularly, reference). However, it seems to me that this rich and thought-provoking paper is equally well read as offering a partial resurrection and vindication of the classical empiricist program in philosophy of mind. Empiricists – such as the early Russell (1912/1998; 1918/1985), Wittgenstein (1922), and Carnap (1929/1967; 1939/1949), or before them, Locke (1689/1975), Berkeley (1710/1962), and Hume (1740/1978), argued that all thoughts were in some way derived from "sense impressions" or "sense data," that is, mental representations of the sensible properties of particular objects. Although they differed as to how they expressed the contents of sense data in natural language ("This is red," "There is a red thing," "Lo! a red thing," "Red here now"), it's pretty clear that what the empiricists had in mind when they talked of sense impressions or sense data were thoughts with precisely the logical form Hurford discusses: A sense datum was a thought to the effect that ($\exists x$)*PREDICATE*(x), where the variable x ranges over objects of sensation and the predicates represent sensible properties of those objects.

The empiricist program in philosophy of mind more or less died out in the 1950s, largely because empiricists made a big mistake about what to count as an "object of sensation" or a "sensible property of objects." For largely nonempirical, epistemological reasons, they denied that sense data could represent *physical* objects and/or *physical* properties and thus hit the wall that ended their research program. A left turn would entail a descent into full-blown phenomenalism: if sense data can't represent physical objects and properties, then things in the world can't *be* physical objects with physical properties – this follows from the facts that (1) we clearly *have* thoughts about things in the world and (2) all thoughts are derived from sense data. A right turn, however, led into a hideous maze of semantic dead ends. In the belief (or hope?) that our thoughts can indeed represent the objective and independent world around us, twentieth-century theorists like Russell, the early Wittgenstein, and Carnap attempted to show how higher-level representations of physical objects and their properties could be reduced to (or constructed out of) sensory primitives. This attempt, notoriously, was a failure, and most philosophers now have abandoned any attempt to explain how our "higher-level" thoughts about tables, dogs, and quarks could be derived from the kinds of information that gets into our minds via our senses. Empiricism died a death, and rationalism – the idea that there's stuff in our minds that *doesn't* come from the senses (I guess it must be innate!) – emerged as the dominant theoretical orientation throughout the cognitive sciences.

In arguing that what initially gets represented in the brain as a result of vision (and, perhaps, audition) is a thought of the form, "THERE'S AN *F!,*" Hurford is in effect arguing that the dead guys were right after all. First, sense data are, as one says, "psychologically real" and do indeed play a central role in the etiology of our thoughts. That is, the brain does represent the *thereness* (or *whereness*) and *whatness* of things, and it does so very early on in the process of thought formation. Second, the old guys were right in thinking that higher-order or more complex thoughts are *derived from* the primitive Hurfordian representations. Turning the Berkeleyan argument on its head: We do have thoughts about independently existing physical objects, and these thoughts do derive from sense data – that is, representations of the form ($\exists x$)*PREDICATE*(x). So, although it's probably not true that thoughts about tables or cats are *constructed out of* or *copies of* sense data (as Carnap or Hume might have hoped), there must be a way to make an empiricist semantics work: The derivation of thoughts from sense data is "psychologically real" as well.

Here, it seems to me, is where the real philosophical interest of Hurford's proposal lies. To pursue the empiricist program further, however, we need answers to two questions. First, what do the variables and predicates of *PREDICATE*(x) range over? Second, what are the processes of derivation, abstraction, inference, binding, (what to call them?) by which these initial representations are manipulated or transformed into the kinds of thoughts expressed by our propositional attitude ascriptions? Hurford has little of substance to say about the second question, but his survey of neuroscientific treatments of the first question is suggestive of the direction in which our semantics for thought needs to move. By making clear that the brain's most primitive predicates seem to carry both "objective" information about objects' spatial location and physical features and "subjective" (egocentric or action-guiding) information, Hurford's discussion in section 4 indicates that their contents are almost certainly quite different *both* from those postulated by the classical empiricists (RED, etc.) *and* from those that are typically lexicalized in natural languages (BOX, LINE, CAT, etc.). This suggests that it may have been a lack of an adequate theoretical vocabulary, more than anything else, that has hindered the empiricist semantic program! Classical empiricists misidentified the contents of sense data – and hence misconceived the nature of the "derivation" process – because they sought to express the "primitiveness" of sense data via "primitive" natural language predicates ("Red," etc.). And because modern investigations into the semantics of thought (whether philosophical or neuroscientific) have had to use horribly complex agglomerates of natural and artificial linguistic forms to express the contents of our most primitive thoughts, they have been taken – wrongly – to be unempiricist in tenor. A new theoretical language for the semantics of thought is needed, one that more closely mirrors the language spoken by the brain. Hurford's paper is an important step in this direction.

Object recognition is not predication

Jean-Louis Dessalles[a] and Laleh Ghadakpour[b]

[a]*ParisTech-ENST, F-75013 Paris, France;* [b]*Ecole Polytechnique, 75005 Paris, France.* **dessalles@enst.fr laleh@infres.enst.fr**
http://www.enst.fr/~jld http://www.enst.fr/~ghadakpo

Abstract: Predicates involved in language and reasoning are claimed to radically differ from categories applied to objects. Human predicates are the cognitive result of a *contrast* between perceived objects. Object recognition alone cannot generate such operations as modification and explicit negation. The mechanism studied by Hurford constitutes at best an evolutionary prerequisite of human predication ability.

Jim Hurford's claim is an impressive attempt to ground human distinctive cognitive abilities like logical reasoning and language in mammalian brain anatomy. His claim is conceptually important to help us understand how a dual where-and-what processing, leading to object recognition, may be a likely prerequisite of human predication. However, the claim that object recognition and predication are similar by nature – differing only in degree – is too difficult to accept, for two groups of reasons.

The first objection is a general critique that can be addressed by any gradualist account of phylogenetic descent. Modern evolutionary theory emphasizes that species are most of the time in equilibrium, and that they qualitatively differ by clear-cut characteristics (Gould & Eldredge 1977). This view is widely confirmed by computer simulations based on genetic algorithms, which show that evolutionary processes are rapid and produce local optima (Dessalles 1996; Holland 1992). One characteristic that our species has in proper is the cognitive ability to manipulate predicates through logical reasoning and to express them through language. The object recognition behaviour shown by mammals is not expected to be either equivalent to, or even to be a draft of, this human ability. One further argument along this line is provided by successful attempts to evolve syntactic language in populations of artificial individuals, as soon as they are granted with some predicate-argument semantics (Batali 2002; Kirby 2000). The fact that other primates seem unable to master syntactic symbolic expressions casts doubt on predicates being available to them.

The second objection against equating object-recognition abilities with (even simple) predication comes from the fact that the underlying cognitive processes are qualitatively different. Jim Hurford restrains the cognitive role of predication to categorisation, "the mental events involved when a human attends to an object in the world and classifies it perceptually as satisfying the predicate in question" (target article, sect. 2). When a perceived object is categorised as an apple, many perceptual features are involved in the recognition: aspects of the shape, colours, textures, presence of two characteristic extremities, and so forth. This ends up, according to Hurford's account, with a predicate like *APPLE(x).* Let us call this process, based on mere object recognition, *R-predication.* Contrary to Hurford's account, we claim that R-predication is qualitatively different from those cognitive processes involved in language that logic represents with predicates. Let us call the latter *C-predication.*

Available models of categorisation, and thus of R-predication, are holistic. Neural networks or standard statistical devices rely on the maximum number of common features that can be found between the object to be recognised and known classes, exemplars, or prototypes. The difficulty, addressed, for example, in conceptual clustering techniques (Michalski & Stepp 1983), is precisely to extract short explicit descriptions for classes and objects. C-predication radically differs by showing nonholistic features: It isolates one *explicit* property from the context. In Ghadakpour (2003), this process is described as resulting from a *contrast operator* (hence *C,* meaning contrast, in C-predication). We are able, without any training, to contrast any object with another resembling object or with its known prototype. This allows us to characterise a perceived object as a *blue apple* or as a *big apple.* Even if the remarkable ability to form prototypes and to see global resemblance between two objects is well within the reach of any mammalian brain, there is little reason to believe that we share with other animals the general ability to *isolate* relevant distinctive properties. Let us mention two reasons.

First, our ability to modify names, like in *a big flea,* has little to do with the co-occurrence of a general property and a location. As Hurford rightfully remarks, the adjective *big* has to be understood here in the flea context (Kamp & Partee 1995). A writing like *BIG*(x) *& FLEA*(X) cannot represent this contextual effect, and it hides the proper cognitive processing. In our own account, both the adjective *big* and the scale on which it is interpreted are provided by the contrast operator. The perceived object contrasts with the prototype of flea by its size; the scale (millimetres vs. metres or light-years) comes from the standard deviation of the prototype; the position of the perceived object on the scale is given by the magnitude of the contrast on this scale. Holistic object recognition (R-predication) does not offer the means to pick a relevant axis and an appropriate scale, so adjective-name modification is strictly beyond its reach.

The second reason why C-predication radically differs from R-predication is that the former can systematically lead to explicit negation. When performing explicit negation, one has to contrast the perceived object with some prototype and to exclude the former from the latter. You may say *this is not a flea, because it is too big*, or *this is not a star, because it is too big.* R-predication alone, because it is holistic, cannot offer such explicit negation. It can only refuse class membership by measuring a holistic distance and comparing it to some threshold. Yet, every human being has the ability to perform explicit negation on any domain without any specific training. This endowment, which underlies the argumentative use of language, is a consequence of our ability to contrast perceptions and form C-predicates. We suggested elsewhere why this ability can be considered as one of the main cognitive differences that distinguishes *homo sapiens* (Dessalles 2000). It offered a new adaptation, namely, the possibility to detect lies. By contrasting one's own perception with the liar's report, one may not only disbelieve the report, but also offer an explicit reason why the report should be rejected. Similarly, because they can be systematically negated, the predicates used in logical accounts of human thinking are C-predicates, not R-predicates. The fact that C-predicates can be used to make membership explicit (*this is a galaxy, because it is big*) may explain why they are mistakenly supposed to be necessary for any categorisation, and hence imprudently granted to animals.

The scope of Hurford's argument is thus more limited than anticipated, because it cannot be extended to "genuine" predication, what we called C-predication. The author's insight and the comprehensive line of argument that he draws from it could yet be extended in some way to C-predication. Our ability to "locate" an object on the axis provided by the contrast operator may be an evolutionary derivative of the fundamental ability to handle location and property separately. To say that this apple is big or that it is bigger than that apple, we must assign positions to objects, not in physical space but on the contrast axis, which may be, in some cases, quite abstract.

Representational limitations of the one-place predicate

Peter F. Dominey
Institut des Sciences Cognitives, 69675 BRON Cedex, France.
dominey@isc.cnrs.fr http://www.isc.cnrs.fr

Abstract: In the context of Hurford's claim that "some feature of language structure maps onto a feature of primitive mental representations," I will argue that Hurford's focus on 1-place predicates as the basis of the "mental representations of situations in the world" is problematic, particularly with respect to spatiotemporal events. A solution is proposed.

Hurford's claim that "some feature of language structure maps onto a feature of primitive mental representations" (target article, sect. 1) is clearly on the right track. However, I will argue that Hurford's focus on 1-place predicates as the basis of the "mental representations of situations in the world" is problematic. Specifically, I will propose that a more appropriate representation is based on the structure of perceptual events that are functionally and behaviorally relevant to the nonlinguistic individual. Such events would include physical contact, transfer of possession of objects, and the like, that inherently consist of multiple-argument predicates. In developing this argument I will exploit Hurford's requirement that the characterization of an appropriate representation should include: (1) "a plausible bridge between such representation and the structure of language" (sect. 1), (2) a characterization of "primitive mental representation" independent of language itself, and finally (3) a plausible story for the neural basis of the representation.

Mapping to language. With respect to the bridge between the representation and the structure of language, Hurford argues that "very little of the rich structure of modern languages directly mirrors any mental structure in pre-existing language" (sect. 1.1). He further states that in contrast to the morphosyntactic complexity of language, the syntax of logical form is very simple. These comments reveal the shortcoming of $P(x)$ as the representation – it is too simple. Indeed, it seems that by focusing on a representation that is appropriate for logic, Hurford steers off the course of a behaviorally useful representational schema. Nonhuman primates likely have quite rich representations of events, their temporal structure, the individuals involved, and so forth. Constructing such representations in a neural first order predicate logic would be difficult. Indeed, the difficulty of the mapping is revealed by the quantity of effort expended in developing a theoretical basis for mapping logic to language and the meanings that can be expressed in language (e.g., Kamp & Reyle 1993; Montague 1970; 1973; Parsons 1990).

I suggest that although 1-place predicates are certainly useful for representing object properties, they are inappropriate for (and do not extend in a straightforward manner to) event representations. Imagine instead that the prelinguistic representation was based on the perceptual structure of events, with ordered predicates yielding a structure something like "event(agent, object, recipient)." In this case the mapping from the mental representation to language becomes more interesting and more iconic. The distinct ordered predicates in the event representation take on specific thematic roles that are iconicly reflected in regularities in word ordering and/or morphosyntactic and closed class structure in a cross-linguistic manner. In the example, "A man bites a dog," the representation: bite(e), man(x), dog(y), agent(x), patient(y) appears arbitrary, unordered, and less informative than bite(man, dog), in which the relations between the event and the constituent thematic roles (agent and patient) are encoded in the representation. I would thus propose that the capability to represent 1-place predicates does not extend in a useful manner to ($n > 1$)-place predicates for representing meaning.

Characterization of the primitive mental representation. Having made this claim, one is obliged to demonstrate the psychological validity of ($n > 1$)-place predicates independent of language. I will approach this from the perspective of (1) observations from developmental psychology and (2) studies of automatic perceptual analysis.

From the developmental perspective, one of the most salient perceptual primitives (after motion) is contact or collision between two objects (Kotovsky & Baillargeon 1998). Prelingual infants appear to represent collisions in terms of the properties of the "collider" and their influence on the "collidee." This supports (but does not prove) the hypothesis that contact is represented by a 2(or greater)-place predicate in prelingual infants.

But is the n-place predicate computationally tractable? That is, is it reasonable to assume that nonlinguistic beings can construct such representations? I have recently explored this question by developing an automated system that extracts meaning from online video sequences of events performed by a human experimenter in a simple setup involving manipulation of toy blocks. The objects are recognized and tracked in the video image, and physical contact between two objects is easily detected in terms of a minimum distance threshold. The agent of the contact is then determined as the one of the two participants that has a greater relative velocity toward the other in the contact. In this context, the event types of touch, push, give, and take can be defined as variants or types of contact events (Dominey 2002; 2003). This demonstrates that sensitivity to a simple class of perceptual event (contact) can provide the basis for a multiple ordered predicate representation of event structure. A more general demonstration of how the perception of support, contact, and attachment can be used to learn the lexical semantics of verbs is provided by Siskind (2001). The objective of developing this perceptual scene analysis system was to demonstrate the feasibility of generating meaning in an event(agent, object, recipient) format, based on the perception of physical contact. This was motivated by simulation studies

http://dx.doi.org/10.1017/S0140525X03380070

of language acquisition based on the learning of mappings between grammatical structure and predicate-argument structures (Dominey 2000), that in turn was based on combined modeling and neurophysiological testing of the underlying functions (Dominey et al. 2003).

These and subsequent studies revealed that the complexity of grammatical forms (e.g., relative phrases) corresponds to an analogous complexity in the predicate-argument representational structure. For example, in mapping the grammatical construct, "The block that pushed the triangle touched the circle," onto the representation push(block, triangle), touch(block, circle), we can observe an iconic relation between the relativized structure of the sentence and the meaning representation in which the two events share a common agent: block (Dominey 2003).

With respect to the neural basis of multiple argument predicates for representing events, one possibility can be found in the F5 neuron populations described by Rizzolatti and Arbib (1998), which, when observed together, allow distinct representations for grasp(me, raisin) versus grasp(someone-else, raisin). Thus, access to two distinct populations of these neurons allows a event representation with distinct agent and object coding.

In summary, I want to insist that Hurford's undertaking is quite valid and interesting with respect to the stated goal of investigating the neural basis of predicate-argument structure. Where it fails is in the thesis, "The structures of modern natural languages can be mapped onto these primitive representations." I hope to have argued that the required representations for events (and their description) are more complex than those described by Hurford – and that they cannot be represented by the primitive structure he describes.

ACKNOWLEDGMENTS
This work was supported by the HFSP Organization and the French ACI for Computational and Integrative Neuroscience.

Cognitive structure, logic, and language

Grant Gillett
Otago Bioethics Centre, University of Otago, Dunedin, New Zealand.
grant.gillett@stonebow.otago.ac.nz

Abstract: Philosophical accounts of thought crucially involve an array of abilities to identify general properties or features of the world (corresponding to concepts) and objects that instantiate those general properties. Abilities of both types can be grounded in a naturalistic account of the usefulness of cognitive structures in adaptive behaviour. Language enhances these abilities by multiplying the experience bases giving rise to them and helping to overcome subjective biases.

Hurford's paper signals a potentially important resonance between Fregean truth conditional semantics and neural structure of the type that philosophers like Evans (1982) and Campbell (1995) have begun to explore.

Underpinning this post-Fregean stream of philosophical naturalism is the idea that thought works with a world-picture comprising objects and their properties, and that such a cognitive map of the world is an adaptive achievement for any higher animal. According to this view, the crucial adaptation is a move away from feature-driven stimulus-response patterns, to tracking abilities based on more complex response sets related to object characterisation and recognition. It is a further move for any creature to share information from multiple subject-object encounters so that a rich dossier of information can be used to inform its behaviour in relation to the world in which it lives. A theory of thought and its content of this kind is indicated by the naturalistic tendency of Wittgenstein's later (post-*Tractatus*) writings and allows truly innovative thinking in theoretical cognitive neuroscience.

We can begin with truth conditional semantics according to which the semantic content of a thought or sentence is given by the conditions in which a competent thinker would hold it to be true. For a creature in a world like ours, such conditions would, stereotypically and ideally, concern an object and a predicate: "the cat is moving"; "the poison berry is red"; and so on. Notice that even if the noun phrase in the sentence concerned is complex (e.g., "the poison berry"), it concerns an object that is salient for the creature thinking the thought. The simplest question that any creature could ask the world would be "what is that?" where an object catches the creature's attention, and it must assign properties to the object so as to potentiate a range of appropriate responses. And the information available to the creature is greatly enhanced if objects can be identified and reidentified (under conditions which yield slightly different perceptual features from the first encounter), so that a dossier for objects of the relevant type can be compiled.

It would be a further step to identify numerical individuals (Blackie) rather than categorical types (big monkey) and the computational resources for such a task would be useful wherever the identification and reidentification of individuals had an adaptive pay-off (e.g., where appropriate responses differed among individuals of the same type as in a dominance hierarchy). Human beings could be expected to have an individual identification system good enough to meet the needs of the environment in which their language developed because exactly this condition holds. These thoughts suggest a three-level scheme of the type mooted in Hurford's article but divorced from the implicit Cartesianism of Frege and many empiricists.

Stage 1: Egocentric space and stimulus response pairings based in individual dispositions – a feature-response scheme.

Stage 2: Egocentric space and activity directed by information about objects as loci of salient general properties – the simple PREDICATE(x) scheme.

Stage 3: Activity (cognitive or overt) in a shared world directed on mutually accessible identified objects and open to information from others – PREDICATE(x').

Human beings have language and interpersonal behaviour as a primary mode of adaptation (Trevarthen 2001), so stages 2 and 3 of Hurford's scheme are modified by signs and conventional referring devices for both general features of the world and particular objects in the world. These facilitate and elaborate the response repertoire of any individual so as to give them access to shared dossiers of information about objects and their affordances. If language is used to expand the informational power of an organism, we need to supplement the three-stage scheme by noting the effect of semantic markers (Gillett & McMillan 2000). True concepts and mature conceptions of objects are tied to truth conditions by the normative uses of natural language, so there is a concurrence of semantic content between colinguistic speakers. Thus, early in language learning I might think that a dog is a big black furry thing that bounds around the neighbourhood, but later I discover that dogs include chihuahuas and poodles. Such convergence in categorisation with other competent language users occurs by conversational correction within a colinguistic human group. By noticing this fact, we can, without denying the continuity between human thought and that of higher animals, bring out a point of difference which increases the power of human epistemic activity and in which language plays a central role.

Attention, a prominent theme in Luria's work (cf. Luria 1973), is important in the formation of concepts and conceptions of objects. In both cases the subject must ground the thought concerned on selected aspects of the environment (Gillett 1992). For example, in the *PREDICATE*(x) type thought "that frog is bright orange," "that frog" focuses on and tracks an object, and "bright orange" links a feature of the frog to other stimulus arrays instancing that colour. Kant (1789/1929) said that thinking was "cognition through concepts"(B94), whereby information from an object was linked in two distinct ways to form discursive or semantic content: (1) to general concepts (square, red, dog, baroque), and (2) to other presentations of the same object. The second link could, in the animal case, be mediated by biological abilities to track the object concerned, but in humans this would be vastly enhanced by the language-related practice of naming things or using denoting expressions (such as definite descriptions – "the black cat from next door"). These, as Hurford notes, are fundamentally deictic or indexical and need only be specific enough to work in the context where they are used (thus, they may not have the logical qualities of unique designators). Wittgenstein's version of naturalism implies that such indexical expressions are "as good as it gets" for embodied creatures banging about in a finite domain.

The present account implies that logic is an idealisation of natural language, which is a tool to aid our cognitive activity, which picks out objects and designates their significant properties. As such, logic formalises our thought content as it appears in language (transformed by the conventional devices that are intrinsic to natural languages). Therefore, logic embodies the fundamental difference between two types of element that make up our cognised world – objects and the properties they instance. It is worth emphasising that a cognitive map of our environment that can represent relatively enduring objects and their properties (including typical trajectories through time and space) is likely to be of great value as we try to make use of the opportunities presented by our world.

Word-sentences and an interaction-based account of language evolution

Bipin Indurkhya

Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan. **bipin@cc.tuat.ac.jp**

Abstract: Considerations from an interaction-based approach to the evolution of language and the role of word-sentences therein show that the object-attribute ontology is arrived at a much later stage. Therefore, Hurford's arguments, by focusing on the predicate-argument structure, seem to miss out on most of the interesting aspects of the early stages in language evolution.

Although Hurford presents persuasive arguments to provide a neural grounding for the object-predicate distinction, they nevertheless constitute a big leap across the landscape of language evolution that seems to miss many of its interesting aspects. In my comments here, I attempt to show a glimpse of some of these skipped-over sights.

For example, when Hurford distinguishes his account with the proposal of Rizzolatti and Arbib (1998), he argues that a structure like "grasp-A(raisin)," corresponding to a canonical F5 neuron as proposed by Rizzolatti and Arbib, is really a shorthand for the two-stage process corresponding to "raisin(x)" and grasp-A(x)." The crux of Rizzolatti and Arbib's proposal, however, is to highlight this evolutionary stage of language where the action is inseparably linked to the object: that is, the stage where the cognitive agent can be aware of something as a raisin only by being aware of how to grasp it. It is only in the later stages of evolution that the action and the object become differentiated corresponding to the two-stage process of Hurford.

Indeed, it has long been recognized that the early stages of language evolution are characterized by word-sentences (cf. Aichison 1996; Barrett 1995; Bloom 1973; Cassirer 1955; Langer 1957; Piaget 1962; Wegener 1885/1971; Wray 1998; 2000). In a word-sentence, the predicate and the object – and various other qualifiers – are fused together, so that the child's first words refer to complete situations. It is only through numerous later interactions that the various objects, attributes, and actions occurring in a situation get gradually differentiated, and the child is able to refer to them individually. Even then, the holistic language continues to play a major communicative role in adult language (Wray & Perkins 2000).

This long evolutionary road leading to the predicate-argument structure can be further highlighted by considering another example mentioned by Hurford: in going from "If it moves, grab it" to "If it catches your attention, inspect it carefully and figure out what to do with it." Along the way, there may be a stage of "If it moves, I can grab it," where the action is contemplated without being carried out. Then, once the action, *grab-it,* and the object, *it-moves,* are differentiated, there may be a stage where each of these schemas is exercised to increase familiarity with it. That is, *grab-it* action is applied to various objects, so that everything becomes something to be grabbed. In this process, the *grab-it* action may become more refined and more detached from the original object. Similarly, other actions are applied to the *it-moves* object, and from these interactions a more comprehensive representation corresponding to the *moving-object* may result. (See Piaget [1953; 1962] for an account of action-oriented interaction.)

This interaction-based approach to language evolution implicitly recognizes the role of a top-down mechanism. The scheme shown in Figure 2 by Hurford suggests a bottom-up mechanism, in which it is the stimuli from the object that evokes the appropriate predicates, attributes, or actions, which may reflect the original affinity of the action and the object, or a habitual association. In the top-down mechanism, however, an action is applied to an unrelated object, as, for example, when a child tries to grab whatever objects he or she may come across. In assimilating an object to an action, novel features of the objects may be discovered (Indurkhya 1998). For example, a child executing the action "grabbing and bringing to the mouth to suck" may notice that the object is visually interesting, too. This, in turn, may lead to the related action "grabbing and bringing in front of the eyes to look at" (Piaget 1953).

Incidentally, it has been proposed that this top-down mechanism forms the cornerstone of metaphors, because it determines the results of the interaction when a predicate is metaphorically applied to an object (Indurkhya 1992). It has even been suggested that the process of noticing perceptual similarities, and thereby generating perceptual metaphors, starts working when the predicates and objects are still in a fused state and can actually underlie the discovery of relations and attributes that hastens their separation (Piaget 1962). In other words, it is in comparing two holistic situations, and finding that they have some features in common, that the independent status of an object or attribute originates. Notice that this account assumes an evolution of memory that is orthogonal to the language evolution in that the cognitive agent, to carry out the comparison, is able to memorize and recall holistic situations before the objects and attributes acquire an autonomous status. (This is in contrast with the view implicit in Hurford's account, where the memory module presupposes object-attribute ontology.)

Finally, it is only at a much, much later stage that a cognitive agent gets to the point where an object is "inspected carefully to figure out what to do with it." Notice that in "inspecting carefully," all those actions that are habitually associated with the object may be activated via the bottom-up mechanism. At the same time, in trying to "figure out what to do with it," actions motivated by the goals and the desires of the agent may be activated via the top-down mechanism. A neural account of any of these steps, if possible, would be quite fascinating.

To conclude, I feel that Hurford's arguments to provide a neural grounding for the object-predicate distinction are a bit like speeding to one's destination in such a hurry as to miss most of the interesting scenery along the way. However, it is the scenery along the way, some glimpses of which I have tried to provide in this commentary, which may well provide the most illuminating insights into the evolution of language and the neural correlates of its different stages.

Predicates as cantilevers for the bridge between perception and knowledge

Gregory V. Jones
Department of Psychology, University of Warwick, Coventry, United Kingdom.
G.V.Jones@warwick.ac.uk
http://www.warwick.ac.uk/fac/sci/Psychology/staff/academic.html#GJ

Abstract: The predicate-argument approach, focused on perception, is compared with the ease-of-predication (or predicability) approach, focused on encyclopedic knowledge. The latter offers functional prediction and implementation in connectionist models. However, the two approaches characterise predicates in different ways. They thus resemble predicational cantilevers built out from opposite sides of cognition, with a gap that is yet to be bridged.

Summarising his approach, Hurford notes in his antepenultimate sentence that "It remains to provide an explanation for the typical structure of modern languages, organized around the Noun/Verb dichotomy" (sect. 6.3), and hints in his penultimate sentence that the explanation may turn out to be exemplified by the various properties of being a dog. This provides an irresistible cue to highlight an alternative predicational account that encompasses not only the theoretical dichotomy but also, strangely enough, the canine properties. This is the ease-of-predication (or predicability) approach (e.g., Jones 1985; 2002).

Whereas Hurford's predicate-argument approach focuses on the relation between predication and perception, the ease-of-predication approach focuses on the relation between predication and encyclopedic knowledge. The ease-of-predication approach makes detailed predictions at a behavioural level by virtue of its characterisation of differences among words in their psychological potential. Normative values of a word's predicability are established on the basis of people's judgements regarding the ease of making statements about the word's referent. The standard instructions (see de Mornay Davies & Funnell 2000; Jones 1985) include the following:

> As an example, the word "dog" would probably be judged as very easy to make simple factual statements about, because it can readily be put into statements such as the following:
> A dog is a type of animal.
> A dog often lives in a kennel.
> A dog barks when angry.
> A dog can be pedigree or mongrel.
> A dog has four legs.
> A dog is called a puppy when young.
> A dog wags its tail when pleased.
> A dog can be as small as a chihauhau.
> A dog can be as big as a St. Bernard.
> A dog sometimes chases a cat. (Jones 1985, p. 6)

The predicates in the preceding illustration may at first sight seem sufficiently varied to make a logician blench, but the variety is intentional because the purpose is to index the ease of access to the whole range of encyclopedic knowledge, unconfined by any particular procrustean formalism. Thus, in logic the examples may be viewed as encompassing predicates that are both simple (e.g., is a type of animal) and relational (e.g., can be as big as a St. Bernard), and also as ranging over the alethic modalities of both necessity (e.g., has four legs) and possibility (e.g., often lives in a kennel). Despite this diversity, when an ease-of-predication judgement for each of a set of words has been established in this way, it turns out to be predictive of several other properties of the word, including the likelihood that it is successfully read by people with deep dyslexia and, for people in general, the likelihood that it successfully retrieves information from memory and that it evokes the experience of mental imagery (e.g., Jones 2002; Williams et al. 1999). Furthermore, the method of assessment of ease of predication is not confined to self-report judgements. Similar results are obtained when assessment is made instead in terms of the response time taken to produce such statements behaviorally (e.g., Jones 2002; Williams at al. 1999), with a high correlation (negative in sign, of course) between words' predication times and predicability judgements.

Are the predicate-argument and ease-of-predication approaches compatible? Having been cantilevered out from opposite sides of the cognitive landscape – respectively, from perception/action and from knowledge/language – it seems reasonable to hope that these two stretches of predicational bridge can be made to join up in the middle. However, there is a problem in the shape of their differing conceptions of predication. In the case of the predicate-argument approach, the interpretation is derived from first-order predicate logic. In the case of the ease-of-predication approach, the interpretation derives most naturally from the linguistic concept of predication as comment on a topic: "The most general characterization of predicative constructions is suggested by the terms 'topic' and 'comment' for their [immediate constituents]: the speaker announces a topic and then says something about it" (Hockett 1958, p. 201). It is rather as though one half of the bridge were designed in metres and the other in yards. What are the incentives for attempting to overcome the difference in gauge?

From the point of view of the ease-of-predication approach, the theoretical parsimony of the predicate-argument approach offers the prospect ultimately of an elegant reductionism. From the point of view of the predicate-argument approach, the most obvious attraction of the ease-of-predication approach is its compatibility with the noun/verb dichotomy – a dichotomy to which Hurford noted that his present formulation appeared to pay little recognition. However, the ease-of-predication approach also offers access to a range of explanatory possibilities at both the functional and neural levels. At one level, it has been shown to lend itself naturally to implementation in the form of artificial neural networks (e.g., Hinton et al. 1993; Plaut 1999). Thus, the influential connectionist model of deep dyslexia proposed by Plaut and Shallice (1993) relies on semantic features distributed in accordance with the ease-of-predication approach, employing only two features for "past" ("has duration" and "refers to a previous time") but 16 features for "post" (e.g., "found on farms" and "used for games or recreation").

At another level, the ease-of-predication approach appears to offer satisfactory functional explanations for findings that have previously been interpreted in terms of mental imagery (e.g., Paivio 1971; 1983). If, for example, either "post" or "past" were now to be used as a retrieval cue to assist the reader in remembering the other, associated word, then previous experimental evidence would suggest that "post," the word with higher predicability – and imageability – would be more likely to succeed in retrieving "past" than vice versa; a finding that can readily be explained in terms of the greater availability of predicational routes from "post," but is difficult to account for in terms of imagery as such. It would be valuable if it were to prove possible to complete a predicational bridge between such findings and Hurford's characterisation of the elements of perception.

Do sensorimotor processes have reflexes in sentence syntax as well as sentence semantics?

Alistair Knott
Department of Computer Science, University of Otago, Dunedin, New Zealand. **alik@cs.otago.ac.nz**
http://www.cs.otago.ac.nz/staffpriv/alik/

Abstract: Predicate logic has proved a very useful tool for the expression of theories of natural language semantics. Hurford's suggestion that predicate–argument structures mirror certain properties of the human sensorimotor architecture can be seen as an explanation of why this is so. Although I support this view, I think that the correspondences that Hurford draws between linguistic and sensorimotor structures not only involve nat-

http://dx.doi.org/10.1017/S0140525X03380070

ural language semantics, but include some elements of natural language syntax as well.

The connections that Hurford draws between structures in predicate logic and the architecture of the human sensorimotor system are very compelling. They provide a strong argument for the idea that the meaning of a sentence consists (at least in some part) in a fairly direct evocation of sensorimotor processes. This idea is already quite widespread in cognitive linguistics – Hurford's main contribution is to ground it very solidly in contemporary models of neuroscience and of linguistic structure. Nonetheless, the more detail that can be recruited in formulating the hypothesis that natural language semantics is grounded in sensorimotor structures, the more convincing it will be. For me, the most exciting contribution of the paper is that it paves the way for new research methodologies in both linguistics and sensorimotor psychology, to generate and test more detailed versions of the hypothesis.

My own interest is in extending Hurford's hypothesis to cover certain aspects of natural language syntax as well as natural language semantics. Hurford (2002) has cautioned against this, arguing that many aspects of syntax will be recalcitrant to this kind of sensorimotor grounding. However, I think that his position in the current paper requires some revision of his earlier argument.

For one thing, Hurford seems to have revised his conception of mental representations since his 2002 paper. In his earlier article, mental representations are emphasised as having no temporal structure: "all parts of the representation of a remembered event are simultaneously present to the mind." In his current paper, however, Hurford is at pains to emphasise the sequential nature of mental representations. The representation underlying a predicate-argument structure is clearly characterised as a sequence of two mental processes: first a direction of attention and then (with implicit reference to this) an action of classification. Hurford's references to Ballard et al. (1995; 1997) reinforce this view: for these authors, mental representations are very largely defined in relation to an ongoing sequence of rapid directions of attention. There is certainly a nontemporal aspect to mental representations, too, on Ballard et al.'s view and in Hurford's current article, because details of the most recently attended-to objects are held in a short-term store, to facilitate reattention. But this is certainly not to say that mental events have no temporal structure.

If mental representations involve sequences of directions of attention, one possible role for syntax is in capturing this temporal structure. I will begin by considering an aspect of syntax that Hurford (2002) singles out as being unlikely to reflect prelinguistic mental representations: the existence of a VP constituent, which groups the verb of a sentence with its grammatical object but excludes the subject. For Hurford (2002), "it seems unlikely that the structure of prelinguistic thought included a VP-like constituent which bracketed a 2-place predicate with one of its arguments, but not the other." But I think the dynamic conception of mental representations in Hurford's current article opens up precisely this possibility.

One of the syntactic motivations for keeping the subject separate from the other arguments of the verb is the possibility of "expletive" subjects; for example, the subject of an existential sentence like, "There is a cup on the table." According to a reasonably generic version of government-and-binding theory, verbs create ("project") syntactic positions for their semantic arguments to fill. The VP is basically defined as the ensemble of these positions. However, subject [Spec, IP] position is not projected by the verb; there is a separate requirement that all sentences must have a subject position. Further, independent grammatical principles allow or require one argument of the verb to "move" to subject position in some circumstances. In the sentence just given, no such movement has taken place; "there" is a phonological expression of an empty subject position.

I suggest that the empty subject in an existential sentence can be thought of as denoting a bare action of attention, and that the constituents in the VP can be seen as denoting psychological processes that are deictically referred to this initial action of attention. For example, the VP-internal NP "a cup" can be thought of as contributing a pure operation of the classification system, occurring in the context set up by attentional capture. In a sentence with a transitive VP (for example, "grabbed the cup"), the verb can be seen as contributing an operation of action classification referred to this same attentional context. Passive sentences tell us that actions can be identified without their agents being classified as objects, but this is certainly consistent with evidence from psychology and neuroscience of a modular faculty for "biological motion recognition" that operates independently from recognition of the agent as an object (see, e.g., Grossman et al. 2000; Johansson 1973). Interestingly, biological motion detection does nonetheless require focal attention (Thornton et al. 2002). The picture that emerges is one in which processes denoted by constituents in the VP are all deictically referred to the initial action of attention associated with subject position. In other words, Hurford's deictic conception of perceptual processing does show some promise as the basis for a sensorimotor characterisation of the subject–VP distinction.

One other syntactic phenomenon is worth mentioning, to extend the sensorimotor interpretation of subject and VP just given. Hurford argues in both the papers I am discussing that the distinction between definite and indefinite NPs should be attributed to the communicative role of language, rather than to its role in representing mental processes. But I think his account of sensorimotor structure allows a lot to be said about this distinction, too. Hurford talks extensively about the notion of an object file: a representation in working memory of an object recently attended to, which serves to facilitate reattention to the object during the course of performing a task. It is interesting to consider the different operations involved in creating and accessing these memory representations. Presumably, a new object file is created when an object is first perceptually established – that is, after an episode of attentional capture. There are also operations when an existing object file is used top-down, to help reattend to an object. These two operations seem very suitable as the sensorimotor denotations of "a" and "the," respectively. "The" is a presuppositional construction: To work out the semantic contribution of "the cup," we must look for a salient referent with the property "cup" in the discourse context, and the construction only has a denotation if we find one. "A cup," however, introduces such a referent into the discourse context. Although I agree with Hurford that referring expressions in communicative contexts need to take into account the user's discourse model, maybe it would be fruitful to associate the set of referents in the discourse context with the set of object files. (The fact that only a small number of object files can be attended to *simultaneously* does not rule out the possibility that a larger set can be stored in short-term memory and used to guide some less detailed form of top-down attention.) This sensorimotor construal of definite and indefinite NPs is certainly consistent with my earlier suggestion about the subject–VP distinction. If existential sentences denote episodes of attentional capture, associated with the creation of new object files, then we expect to find indefinite NPs in these contexts and not definite ones, and indeed a sentence like "There is the cup on the table" is ill-formed.

Perceiving and describing motion events

Shulan Lu and Donald R. Franceschetti

Psychology Department and Institute of Intelligent System, The University of Memphis, Memphis, TN 38152. **slu@memphis.edu** **dfrncsch@memphis.edu**

Abstract: According to Hurford, PREDICATE (x) is correlated with deictic object variables during event perception. This claim is inconsistent with some core literature on the perception of motion events. We point out that the perception of events involves the activation of the modal properties and amodal properties of underlying event structure, for which Hurford's target article fails to account.

Arguing that PREDICATE (x) is a schematic representation of object properties and spatial information, Hurford approaches the problem by dissecting the components of the predicate logic and, correspondingly, how attention is partitioned when an event is perceived. For example, variables in FOPL tend to be instantiated by whole objects. Hurford has reviewed the relevant literature and does find evidence supporting this position. This raises a question: How do we perceive and describe events?

Imagine a man living 50,000 years ago. How does he survive? He tracks, follows, and hunts animals. He tries to avoid some animals. He and his people definitely need to have certain variables that convey the features of the animals. More important, he needs to perceive and interpret the motions of the objects around him. How do people describe motion events? Talmy (1975; 2000) suggested that there are the following four semantic components of motion events: (1) the Figure: an object moving or located with respect to another object (the Ground); (2) the Motion: the presence per se of motion or location in the event; (3) the Path: the course followed or the site occupied by the Figure object with respect to the Ground object; and (4) the Ground. Furthermore, there is an external co-event that often bears the relation of Manner or Cause to the motion event. The following is Talmy's own example that demonstrates all four of the semantic components and the notion of co-event (Talmy 2000, p. 26):

Table 1 (Lu & Franceschetti). *From Talmy 2000: Semantic components and co-events*

	Manner	Cause
Motion	The pencil rolled off the table.	The pencil blew off the table.
Location	The pencil lay on the table.	The pencil stuck on the table (after I glued it).

In the preceding example, the pencil is the Figure, and the table is the Ground. Off and on encode the Path (or a site). The verbs in the top row encode the motion, and the verbs in the bottom row encode the location. Besides the states of Motion, the Manner is expressed by rolled and lay, whereas the Cause is expressed by blew and stuck.

Cross-linguistic studies have suggested that these semantic components are part of the conceptual structure despite the fact that different languages have different surface forms (Talmy 2000). For example, English language tends to use prepositions when expressing the Path, whereas Spanish tends to package the Path into the verb. This indicates that the perception and interpretation of motion events is more than the mapping of the predicate logic and the activation of the relevant components in the brain. The perception of motion involves the segmentation of motion.

How do we partition the stream of motion? Zacks and Tversky (2001, p. 10) proposed, "The building blocks of events should be temporal units in which the figure, motion, path, and ground are constant. A change in any of these features of the situation constitutes a new atomic event." For example, as a hunter sees a deer, the hunter quietly approaches the deer. The deer seems to be noticing something. It turns its body and looks around. Very likely the change in the motion of the deer is perceived as an event. The hunter hides himself in the bushes, holding his breath. As the deer resumes its activity, the hunter creeps toward the deer again. This time the deer spots the movement. The deer runs. The hunter chases the deer. The deer suddenly takes a left turn. Here a change in Path occurs. Quickly the deer gets into the forest. Here a change in Ground occurs. The hunter sees that the deer is running, therefore, the hunter accelerates his speed. Zacks and Tversky (2001) suggested that changes in Figure may not lead to a new sub-event. However, the contrary is very much possible, as the following example suggests: The deer runs into its herd, and the herd starts running away from the hunter. In other words, the perception of events involves perceiving the changes and segmenting the stream of information.

Zacks et al. (2001a) identified a network of brain regions tuned to perceptually salient event boundaries under both passive viewing and active segmentation of videotapes of goal-directed everyday activities. The segmentation involves the differentiation of the aspects associated with motion, which could not be accounted for by updating a few variables of object profile. By examining patients who are incapable of perceiving different motions, studies in neuroscience suggest that the perception of motion involves more than the striate visual cortex (Bear et al. 2001).

When the hunter sees the herd, what is his mental representation? According to the predicate-argument structure, his representation is herd (deer). However, we tend to treat the herd itself as an object. The argument here is that the perception of events is more than the output of an object variable system (Hurford's major claim). The perception of events involves the perception of the underlying structure of an event, which is amodal (Gibson & Spelke 1983). For example, viewing motion events could activate the perception of causality (Michotte 1946/1963; Scholl & Nakayama 2002). The hunter throws the flint toward the deer. To capture the deer, the hunter needs to know the angle and velocity of the flying flint. There is evidence showing that people tend to spontaneously construct a hierarchical structure of mundane goal-directed events, such as hunting, during the online perception of the event (Zacks et al. 2001b).

In summary, the perception of events not only involves the perception of the modal properties, but also the amodal underlying structure of the incoming information. This is where the current account falls short.

Message *and* medium: Lowly *and* action-related origins

Peter F. MacNeilage[a] and Barbara L. Davis[b]

[a]*Department of Psychology, University of Texas at Austin, Austin, TX 78712;* [b]*Department of Communication Sciences and Disorders, University of Texas at Austin, Austin, TX 78712.* **macneilage@psy.utexas.edu babs@mail.utexas.edu**

Abstract: Hurford presents a much-needed lowly origins scenario for the evolution of conceptual precursors to lexical items. But more is still needed on *action,* regarding both the *message* level of lexical concepts and the *medium.* We summarize our complementary action-based lowly origins (frame/content) scenario for the vocal auditory *medium* of language, which, like Hurford's scenario, is anchored in a phylogenetically old neurological dichotomy.

When a future history of the understanding of language evolution is written, a major twenty-first century contribution may have been the widespread gut-level realization of the relevance of Darwin's "lowly origins" metaphor: "in his bodily form, man bears the indelible stamp of his lowly origins." (Darwin l871, p. 597) As form is in the service of function, we must necessarily seek a lowly *functional* origin for language. A number of issues have blocked consideration of the deep historical roots of modern language, Including our characteristic anthropocentrism, the Western religious dichotomy between man and animal (endorsed by Descartes), the scientific tradition of considering language as autonomous, and the consequent tendency within the formalist linguistic tradition to ignore the question of prelinguistic evolution altogether. For most people, what could be simple ental precursors are seen as altogether too remote from language. Hurford has the courage to challenge this intuition by proposing a phylogenetically deep and therefore necessarily multistage scenario.

One criticism of Hurford's perspective is that he could have made more of his second major conclusion from research on the neurology of vision – "that much of the visual processing in any organism is inextricably linked with motor systems" (target article, sect. 2.1.1, referring to Milner & Goodale 1995). He does note that Bridgeman et al.'s label of "motor oriented" for the dorsal system might be better than the label *where* (sect. 2.1.1; cf. Bridgeman et al. 1979). He also proposes that this system must play a role in attention and orientation precursors to deixis, which is in turn a precursor to labeling, and thus nominalization. Nevertheless, Hurford's scenario remains too reminiscent of the classical approach to mind as an organ of perception of the world and cogitation, but not the organ of action that the Darwinian theory of natural selection calls for. What is selected is successful use. As Mayr has said, "Behavior (is) the pacemaker of evolutionary change" (Mayr l982, p. 612). There are two issues here. First, Hurford neglects the importance of action beyond deixis at the conceptual or *message* level. Second, if one supposes that relatively early stages of language evolution occurred in a communicative context, the message level must have been interfaced with the *medium,* and for this, an action component was also essential.

Consider the message level. An early origins scenario must get beyond deixis to the fact that *what you do about something* – the action it affords – is part of the mental conceptualization of a large number of environment entities, especially for predators, and the consequences of this must have been important in the evolution of syntax from the very beginning. The importance of action in conceptualization today is clearly indicated by the presence of a brain region that apparently mediates category-specific lexical representations (names) for tools. These representations must necessarily contain information regarding the actions that the tools afford. The brain region – the posterolateral part of the left inferior temporal lobe, and the lateral temporo-occipito-parietal junction (Damasio et al. l996) – is part of the *what* system, for example, a chimpanzee looking at a sapling as a potential termite wand. The subsequent decision must be based on a representation of action affordances. Another lowly origins scenario for the anterior cingulate cortex, a major player in attentional control and in *selection of appropriate action* (Gazzaniga et al. 2002), would be a desirable addition to Hurford's proposal.

Despite this criticism, we are sympathetic to Hurford's proposal because, like him, we also are trying to sketch out a lowly origins approach to the evolution of language. We are trying to do for the *medium* what he has done for the *message.* We also emphasize the likelihood of a multi-stage evolutionary process in contrast to the single-stage scenario for both grammar and phonology favored by Chomsky and many other generative linguists. Furthermore, we, like Hurford, are trying to find a wormhole between linguistics and neuroscience in relating basic properties of language to a well-accepted dichotomy of mechanisms at the neurological level (see Goldberg l985). We see our approach as complementary to Hurford's, and consequently, we will briefly spell it out here.

Our frame/content (F/C) theory is a conception of the evolution of the syllable according to which this basic unit of the vocal–auditory medium has deep phylogenetic roots. (See MacNeilage l998; and also MacNeilage & Davis 2000; 2001). The motor basis for the close/open alternation of the mouth characteristic of the syllable (closed for consonants, open for vowels) is oscillation of the mandible between elevated and depressed configurations. This movement cyclicity is proposed to have evolved for various ingestive processes (chewing, sucking, licking) in early mammals about 200 million years ago. The mandibular cycle then may have been used for visuofacial cyclicities (lipsmacks, tonguesmacks, teeth chatters) as observable today in many higher primates (Redican 1975), before it was paired with phonation in hominids to form the frame for the syllable. Subsequently, the capability of programming the frame with independently controllable content elements (consonants and vowels) arose, in response to selection pressures to increase the message set.

In our view, the frame/content dichotomy has its neural basis in a dichotomy of motor control subsystems present in all primates. One is an "extrinsic" system, a system that in humans allows movements to be influenced by multimodal perceptual input, whether the movements are associated with the vocal or the manual system (Goldberg l985). The other is an "intrinsic" system that is primarily involved in self-generated behavior, whether manual or vocal (Goldberg l985).

A basic neurological fact relevant to the proposed evolutionary origins of the syllable is that mechanisms in ventral motor and premotor cortex (the latter area including part of Broca's area in humans) are specialized for the control of ingestion processes in mammals including ingestive cyclicities (Luschei & Goldberg 1981). A key recent finding in this regard is that actions of both oral ingestion and lipsmacking have claims to be implicated in the evolution of a mirror neuron capability (Rizzolatti & Arbib 1998) in area F5 of monkey ventral premotor cortex, the homologue of Broca's area (Fogassi 2003). This capability may be a precursor to the evolution of the ability of hominids to learn speech.

As suggested by its responsiveness to external input, ventral premotor cortex is part of the extrinsic system. We argue that in modern humans, whereas the ventral premotor cortex plays the main role in the production of segmental content, frame production is primarily the province of the intrinsic system, particularly the supplementary motor area. The latter claim is based mainly on evidence that the SMA is implicated in the involuntary production of strings of identical consonant–vowel syllables in a number of instances of electrical stimulation or irritative lesions of the SMA and instances of apparent disinhibition of SMA output in global aphasics (MacNeilage & Davis 2001).

There are major differences between Hurford's approach to the problem of language evolution and ours, differences that are perhaps inherent in the subject matter. For example, Hurford links the message level of language with logic, while we link phylogeny and ontogeny at the level of the medium, primarily on basic biomechanical grounds. What we have in common is an attempt to establish evolutionary links between basic phenomena in linguistics and cognitive neuroscience, links that could help to reveal what we both believe to be the inevitable lowly origins of language.

ACKNOWLEDGMENT

This paper was prepared with support from research grant No. HD 27733-09 from the Public Health Service.

Arguments in the syntactic straitjacket

Massimo Piattelli-Palmarini and Heidi Harley

Department of Linguistics, University of Arizona, Tucson, AZ 85721-0028.

massimo@u.arizona.edu hharley@u.arizona.edu

http://w3.arizona.edu/~ling/hh/

Abstract: While the search for the neural basis of the language of thought is a laudable enterprise, and the article by Hurford a valiant first attempt, we argue that in investigating the argument structure of natural language it will ultimately prove more fruitful to consider the restrictions forced on the system by its inherently syntactic character.

Although the success of this kind of project is devoutly wished, it is at variance with Hurford's own characterization of it as "reductionist," and we prefer to characterize it as constructively "translationist." In this brief commentary, we intend to explain why, in our opinion, it is only partially successful, and what would constitute success for a better project.

We start by focusing on what is arguably the central problem (CP) in the study of arguments and argument structure: the paucity of their number and the rigidity of their character across the languages of the world. This datum stands in stark contrast with the myriad properties and relations that may be relevant to the most ordinary commerce of living creatures with the surrounding world. We may well expect a living creature to entertain judgments concerning an action or an object that are sensitive to properties such as edibility, appeal, danger, source, rarity, risk, and ease, to name just a few. Yet, the number of thematic roles that may be associated with any single predicate is severely restricted, at most three, possibly four. All other judgments can easily be supplemented linguistically by means of adjunction or conjunction: on Tuesday, with alacrity, and it was made of bamboo, and the warrior handled it deftly, and so on.

Hurford says: "In the view adopted here, a predicate corresponds . . . to a judgement that a creature can make about an object" (sect. 1.2). He duly acknowledges the existence of the CP: "The simple clauses of human languages are constrained to a maximum of about four or five core arguments; indeed, most clauses have fewer than this. Presumably this reflects the structure of the underlying mental propositions" (sect. 4). The explanation offered is crucially in terms of the prelinguistic (perceptual, mnemonic, computational constraints, allegedly supported by neurobiological data).

In the target article, predicates are presented as characteristically derived from propositions. This is basically the Aristotelian conception, which does not withstand serious scrutiny. Hale and Keyser (2002), leading proponents of the "constructionalist" approach to argument structure, have explained the scarcity and the rigidity of arguments in terms of the very limited number of possible syntactic nodes projected by predicates (notably verbal predicates) in hierarchical phrasal trees, and the intrinsic syntactico-semantic value associated with these nodes. If this account is even approximately correct, the explanation of the CP lies in the development of the linguistic apparatus. The constraints are language-internal, hierarchical, and structural, not perceptual, conceptual, mnemonic, computational, or otherwise imposed by some extralinguistic system. They indeed "reflect . . . the structure of the underlying mental propositions," though in a sense rather different than the one suggested by Hurford.

A different theory of argument structure, a "lexicalist" one (Jackendoff is rightly quoted by Hurford) lends itself more to the picture suggested in this paper. In a Jackendovian semantics, a lexical conceptual representation is composed of several irreducible conceptual elements (predicates), each of which may contribute one or more argument positions to the final lexeme. Jackendoff argues that these conceptual primitives owe their existence and character to extralinguistic three-dimensional perceptual representations. However, Jackendoff is no closer to accounting for the CP than Hurford is. The construction of a lexical conceptual structure could in principle involve an arbitrary number of primitives and hence an arbitrary number of arguments. The numerical limitation is not imposed by formal logic, where the expectation for n-ary predicates is that n could equal any natural number, any more than it follows from restrictions imposed by our perceptual system.

Concerning reference, Hurford says,

> Information about the relative location of the objects in a visual scene, or about the properties of those objects, represented in a perceiver's brain, has the same essential quality of "aboutness" . . . that linguists and philosophers identify with the semantics of sentences. Those . . . who have insisted that semantics is a relation between a language and the world, without mediation by a representing mind, have eliminated the essential middleman . . . The vision researchers have got it more right, in speaking of the "semantics" of neural representations, regardless of whether any linguistic utterance is involved. It is on the platform of such neural representations that language can be built. (target article, sect. 5, para. 5)

The truth-functional semantics of natural languages, in the wake of Frege and Tarski, has eliminated the subjectivity of the particular speaker, but not the speaker's tacit knowledge of language, nor internal representations of expressions in the "language of thought" (Fodor 1975; 1987; Fodor & Lepore 2002). Reference is always made to entities "under a description" (Chomsky 1988), that is, via tokens of internal standard representations in Mentalese. Neural states or processes as such have no semantics. They co-vary nomologically and causally with events in the world. Only symbols (bona fide representations) can have a semantics, and representations are descriptions accessed internally by the subject. It's very important to determine how the neural apparatus constrains those representations, but the locus of semantics is in those representations, not in their supporting neuronal states and processes.

A sophisticated neuroscience can ascertain the nature of the "raw" perceptual inputs to the representational apparatus, and the constraints imposed on the relevant mental computations by the hardware that implements them. For example, an intrinsic sensitivity to events as primary percepts (as foreshadowed by Lettvin et al. [1959] and competently refined and updated here by Hurford) is surely evidence in favor of the central role that events play in the semantics of natural languages (Higginbotham 1985; 1989; Schein 1993). The split between neuronal pathways sensitive to locational and intrinsic properties of objects and events can indeed offer the ingredients of an explanation of the chasm between closed- and open-class lexical items. These are, of course, promissory developments, only dimly foreshadowed by present-day neurobiology, in spite of its relentless and awesome progress.

In sum, this argument structure is currently analyzed by two distinct theories, *lexicalism* (Jackendoff 2002; Pustejovsky 1995) and *constructionalism* (Borer, in press, Goldberg 1995; Hale & Keyser 2002). Should the considerations developed by Hurford withstand a revision along the lines proposed here, lexicalism might win some points. In the end, we suspect that the contest between theories will be determined by internal coherence, and the ability to predict and explain relevant linguistic phenomena. Yet, cogent evidence from the basic neurosciences may well play a role. We are not advocating reductionism, but rather a sensitivity to the results of a productive intertranslation between domains. Probably, even Frege and Russell would have welcomed the opportunity to test their semantic theories on a terrain that was still inconceivable in their own time.

No problem for Aristotle's subject and predicate

Guy Politzer

Psychologie Cognitive, Université de Paris VIII, Centre National de la Recherche Scientifique, 93526 Saint-Denis, France. **politzer@univ-paris8.fr**

Abstract: It is argued that, in the traditional subject-predicate sentence, two interpretations of the subject term coexist, one intensional and the other extensional, which explains the superficial difference between the traditional S-P relation and the predication of predicate logic. Data from psychological studies of syllogistic reasoning support the view that the contrast between predicate and argument is carried over to the traditional S-P sentence.

That the predicate-argument structure of modern languages, on the one hand, and of logical formulae, on the other hand, have a precursor in the integration by the brain of two pathways, one to locate an individual object and one to identify properties, seems plausible. This division of labor between predicate and argument is well respected by first-order predicate logic (FOPL). But, at first sight, it would seem to be at variance with the Aristotelian sentence (henceforth, A-sentence), which *superficially* does not require the argument (the subject) to be an individual variable, nor does it require the subject and the predicate to belong to different kinds of terms (i.e., the same term can be the subject of one sentence and the predicate of another). This is what the author calls "the Aristotle problem," to which a simple solution will be proposed.

It will be useful to separate two senses of *predicate:* (1) in its modern, logical sense, a predicate (henceforth, predicate_L) is a function from a singular term to a sentence expressing a proposition about the object to which the singular term refers; and (2) in its traditional, grammatical sense, the predicate (henceforth, predicate_G) is that which is affirmed or denied of the other term, the subject. I believe that the solution to the problem at hand lies in a dual interpretation of the subject term. In his conception of terms, Aristotle took the intensional or the extensional point of view, depending on the domain explored (putting greater emphasis on the former in his theory of the proposition and near exclusive emphasis on the latter in his theory of the syllogism). I suggest that the Aristotelian subject should be treated (by theorists) and can be processed (by speakers) from both points of view. With an intensional reading, the subject functions as a predicate_L (which trivially licenses the occupation of the subject slot and the predicate_G slot by the same term). But with an extensional reading, the subject functions as a class providing generic *individual* terms, which licenses the occupation of the subject place to play the role of an argument in association with the predicate_L. In brief, the traditional A-sentence superficially expresses a higher-order predication, but its generic subject term is of the same logical type as the other arguments to which the predicate_G may apply (*qua* predicate_L): Predication in the A-sentence is understandable on the same grounds as the basic predication, from which it does not differ in nature.

From the inception of categorization, the A-sentence predication could start to develop, taking *generic* individual objects as its subject: "*x* is a predator" (with temporal anteriority) and "*x* is yellow" are conflated into "the predator is yellow." Indeed, it would be uneconomical to formulate, for example, *the x is S and P* when the first predication is temporally or cognitively already established; hence the shorter formulation *the S is P.* Now, in remarking that "FOPL is more distanced from the surface of natural languages" (sect. 6.3, last para.), Hurford makes an understatement: FOPL, at least with standard quantification, fails to capture what the A-sentence captures by using the generic term of the subject class, namely, the anteriority of the first predication (Sx) with regard to the second (Px). For example, in the rendering of *some cat is black,* $(\exists x)\,(Cx\ \&\ Bx)$, C and B are treated on a par, except if one remedies this by using restricted quantification, $(\exists x\!: Cx)(Bx)$, which captures well the prior predication.

There is indirect evidence that for Aristotle the deep structure of sentences like *some S are P* coincides with their expression in FOPL. This is suggested by one of his methods to solve syllogisms, namely *ecthesis.* Consider his proof for Darapti: *all M are P; all M are S; therefore some P are S.* First, extract an individual *x;* now, because both *P* and *S* are predicated of this individual, it follows that one of the two is predicated of the other: *some S are P.* In other words, *some S are P* is proved by the fact that there is an individual that has both properties *P* and *S.* This turns out to be a paraphrase of the FOPL formula $(\exists x)(Sx\ \&\ Px)$. But notice that in this syllogism, *S* and *P* play parallel roles (except in the conclusion where the end term labeled *S* is, by convention, allocated to the subject place), so that none of the terms *S* and *P* takes precedence over the other; this leads to another psychologically acceptable conclusion *some P are S.*

If, as I have argued, the nature of the predication in the A-sentence does not differ from the basic predication in which (in agreement with Hurford's claim) the object attended to (the subject) and the property attributed to it have a different status, it follows that people should be sensitive to this difference in status. Of course, this meets speakers' intuitions of a contrast between subject and predicate_G; but also this ought to have testable consequences: People should be reluctant to exchange the role of the subject and of the predicate_G. This is precisely one of the most robust and remarkable findings in the psychological study of syllogistic reasoning. There are two related phenomena. One is called the *figural effect* (Johnson-Laird & Steedman 1978). Irrespective of the logical status of the syllogisms, people have a tendency to produce a conclusion whose terms (subject and predicate_G) reproduce the role which they have in the premises. Take, for example, the syllogism *some M are P; all S are P,* from which most people fail to recognize that nothing follows; instead, they provide *some* conclusions, most of which are *some S are P,* and only very few *some P are S.* The other phenomenon concerns valid syllogisms. The four most difficult syllogisms to solve (for which the success rate is typically below 15 percent) coincide with the only ones in which both end terms change grammatical role from premise to conclusion (e.g., *no P are M; all M are S; therefore some S are not P*).

Ventral versus dorsal pathway: The source of the semantic object/event and the syntactic noun/verb distinction?

Markus Werning

Department of Philosophy, Heinrich-Heine-University, Düsseldorf, D-40225, Germany. **werning@phil-fak.uni-duesseldorf.de**
http://www.phil-fak.uni-duesseldorf.de/thphil/werning

Abstract: Experimental data suggest that the division between the visual ventral and dorsal pathways may indeed indicate that static and dynamical information is processed separately. Contrary to Hurford, it is suggested that the ventral pathway primarily generates representations of objects, whereas the dorsal pathway produces representations of events. The semantic object/event distinction may relate to the morpho-syntactic noun/verb distinction.

By presuming that expressions of language exhibit certain logical forms and by mapping these forms onto regions of cortex, Hurford tries to reduce conceptual structures to neural events. Although I agree to this two-step method, I recently presented a methodologically similar approach (Werning 2001; 2003a) that arrives at different conclusions with regard to the cortical realization of object and property concepts: (1) A property concept is identified with the column(s) of neurons that code for the property in question. The application to the ventral/dorsal division leads to the conjecture that property concepts are located both in the ventral and the dorsal stream. However, the ventral stream may predominantly host representations for properties of *objects* expressed by objectual predicates (e.g., by adjectives like *red, vertical,* or, in more complex cases, by nouns like *square, circle*), whereas columns in the dorsal stream may predominantly represent properties of *events,* for example, the direction or speed of motion. These are expressed by eventual predicates (e.g., by adverbs like *downward, quickly,* or, if complex, by verbs like *to fall, to move*). (2) Individual (object and event) concepts are identified with synchronous oscillations. (3) One or more property concepts $F_1, \ldots, F_n$ are predicated of an individual concept *x,* that is, $F_1(x)\ \&\ \ldots\ \&\ F_n(x)$, just in case neurons of the columns that code for the properties represented by $F_1, \ldots, F_n$ fire in synchrony, that is, with the same oscillation function. On the basis of these assumptions it was possible to show how cortical structure semantically realizes a predicate language (Werning 2003b).

To highlight the differences to Hurford's proposal, let me first turn to what I call his *object-by-position hypothesis* (OP): An object concept amounts to the representation of the object's position in space. OP expresses a fairly rich notion of what object representations are, because falling under an object concept would be equivalent to having all the positional properties the object concept represents. In other words, being object *x* requires the possession of certain *essential* properties, if only positional ones. My view, in contrast, is less demanding and follows Quine's (1961) dictum: "To be [an individual] is to be the value of a [bound] variable." Accordingly, an individual concept would be the neural

http://dx.doi.org/10.1017/S0140525X03380070

event that binds together neural representations of properties. Following Singer and Grey's (1995) theory of binding by synchrony, individual concepts should be identified with synchronous oscillations.

Hurford conjoins OP with what I call his *position-through-dorsal hypothesis* (PD): The dorsal stream (plus superior colliculus and pulvinar) is predominantly engaged in the processing of positional information, and particularly not in the processing of property information. A rival view holds that the dorsal stream primarily processes motional information (Wurtz & Kandel 2000). According to Merigan and Maunsell (1993), the dorsal stream – involving, inter alia, the thick stripes, the middle temporal (MT), medial superior temporal, lateral parietal and ventral parietal areas, and area 7a – receives input predominantly from the magno pathway via lateral geniculate nucleus. Recall that magno cells have excellent dynamic (temporal) resolution, whereas the parvo cells contributing mainly to the ventral stream have much better static (spatial) resolution. Furthermore, MT of the dorsal stream seems to be paradigmatically involved in motion processing. The dorsal pathway can thus be regarded as carrying mainly dynamic, that is, eventual information (prototypically motion), whereas the ventral stream seems to be preoccupied with static, that is, objectual information (prototypically, color and form). It thus remains an empirically unsettled issue whether the ventral/dorsal division corresponds to a distinction between property ("what") information and positional ("where") information, or to one between objectual and eventual information. Wurtz and Kandel (2000) review a large amount of data from lesions in humans and monkeys that support the second option.

Even if one were to accept OP and PD, it would be rash to conclude that object concepts are delivered exclusively by the dorsal stream. For, if property concepts are processed by the ventral stream, what then is the mechanism of predication, that is, the mechanism of binding an object concept to a property concept? Hurford gives no answer. Theorists who identify individual (object and event) concepts with oscillation functions, in contrast, have shown in detail how the neurons of one column can be modeled as oscillators so that the Gestalt principles are honored, according to which neighboring elements with similar properties are likely to belong to one and the same individual (Maye 2002; Schillen & König 1994; Werning 2003b). According to this view, an individual concept is generated within hyper-columns by synchronizing and desynchronizing connections. Because columns serve as property concepts, there is no anatomical separation between the processing of property and individual concepts.

Summing up the critical arguments, one may contrast Hurford's view with an alternative hypothesis: Property concepts and individual concepts alike are processed in the ventral and dorsal stream. However, the dorsal stream is predominantly occupied with the representation of events, which are dynamic in nature. It hosts concepts of eventual properties and generates individual event concepts. The ventral stream, on the other hand, tends to produce representations of objects. It hosts objectual property concepts, which are static in nature, and generates individual object concepts. This ontological division of objects and events reflects a structure that is well known from the logical analysis of language and thought (Varzi & Pianesi 2000). A sentence of the form, "A red circle is slumping," has to be analyzed by quantification over an object (the red circle) and an event (the slumping): $(\exists x \exists e)(\mathrm{RED}(x)\ \&\ \mathrm{CIRCLE}(x)\ \&\ \mathrm{SLUMPING}(e)\ \&\ \mathrm{AGENT\_OF}(x,e))$. Hence, the mental representation expressed by the sentence consists of two objectual property concepts (RED and CIRCLE), one individual object concept (x), one eventual property concept (SLUMPING) and one individual event concept (e). According to the alternative hypothesis, the neural realization of RED and CIRCLE should be columns of neurons in the ventral stream (e.g., V4). There should be an oscillation among them that corresponds to the object concept x. Furthermore, the property concept SLUMPING is expected to be realized by columns of neurons in the dorsal stream (e.g., MT), and those neurons are predicted to oscillate synchronously in a way described by the oscillation function that corresponds to e.

The alternative theory would, moreover, allow us to aim at a neurobiologically founded explanation of the origin of the morpho-syntactic noun/verb dichotomy. Although its universality has been disputed, there seems to be rich evidence that it holds (Croft 2000; Mithun 2000). Nouns and their modifiers – adjectives – prototypically denote objects and their properties, whereas verbs and their modifiers – adverbs – prototypically refer to events and their properties. Because the alternative hypothesis suggests that the semantic object/event distinction correlates with the ventral/dorsal division, one might conjecture that this division, at least in evolutionary terms, is the origin of the noun/verb distinction.

The neural representation of spatial predicate-argument structures in sign language

Bencie Woll

Department of Language and Communication Science, City University, London, United Kingdom. **b.woll@city.ac.uk** **http://www.city.ac.uk/lcs**

Abstract: Evidence from studies of the processing of topographic and classifier constructions in sign language sentences provides a model of how a mental scene description can be represented linguistically, but it also raises questions about how this can be related to spatial linguistic descriptions in spoken languages and their processing. This in turn provides insights into models of the evolution of language.

Hurford's target article proposes a "wormhole" between formal logic and empirical neuroscience, identifying *PREDICATE(x)* as a schematic representation of the brain's integration of the location of an arbitrary referent object, mapped in parietal cortex, with the analysis of the properties of that referent by other systems. A single point will be raised here for consideration in relation to this proposal. With this model, it might be expected that the parietal lobes would be involved in linguistic comprehension tasks, especially those that demand spatial representational resources. In nonlinguistic contexts, a very wide range of spatial functions is associated with parietal lobe function (see Culham & Kanwisher 2001 for a review). However, even when space is referred to in spoken language, there is little evidence that these parietal systems, specialised for spatial processing, are specifically activated.

Although parietal regions may be involved in tasks such as solving spatial syllogisms (Carpenter et al. 1999) or generation of spatial prepositions in response to visual images (Damasio et al. 2001), this does not appear to be mandatory (Goel et al. 1998; Reichle et al. 2000). Indeed, there is evidence that the parietal involvement in the Damasio et al. study may arise from the processing of the visual image and not from the linguistic task itself. Because the claim is made by Hurford that mapping of "scenes" by the parietal cortices underlies the subsequent creation of linguistic structures, the absence of any parietal involvement in language processing needs to be explained.

In relation to this point, data from sign language research is of interest. Although the sign languages of contemporary deaf communities do not provide direct evidence relating to the evolution of human language, because they have arisen in humans with "language-ready" brains, they do provide insight into what an earlier "wormhole" might have looked like.

In sign languages, space serves several functions. All signing occurs in "sign space," an area in front of the signer. This space may be regarded in different ways: From a phonological perspective, it serves simply as a region for the execution of signs. At a higher level, entirely abstract sentence meanings can be represented spatially. In the BSL (British Sign Language) translation of the sentence, "Knowledge influences belief," one location in the space in front of the signer is assigned to "knowledge," a second location to "belief," and the verb "influence" moves from the location assigned to "knowledge" towards that assigned to "belief." Such sentences may be regarded as exemplifying a *referential* use of space, in which spatial relations are used to differentiate grammatical classes and semantic roles. In such sentences, and even in less abstract BSL examples, such as "The woman keeps hitting the man," the locations of events in sign space do not represent and are not constrained by "real-life" spatial relations. However, in addition to these functions, BSL sentences can be constructed topographically. In topographic sentences, "the linguistic conventions used in this spatial mapping specify the position of objects in a highly geometric and nonarbitrary fashion by situating certain sign forms (e.g., classifiers) in space such that they maintain the topographic relations of the world-space being described" (Emmorey et al. 1995, pp. 43–44).

Because of this link between real-world spatial representations and language in such constructions, there has been recent interest in how sign languages may make use of cortical systems specialised for spatial processing and how this may differ essentially from what is found in spoken language (Campbell & Woll 2003). Two recent functional imaging studies of sign language processing (Emmorey et al. 2002; MacSweeney et al. 2002) have cast some light on the question.

MacSweeney et al. (2002) used fMRI to explore the extent to which increasing the topographic processing demands of BSL signed sentences was reflected in the differential recruitment of parietal regions. Enhanced activation was observed in left inferior and superior parietal lobules during processing of topographic BSL sentences (e.g., "The pen is to the left of the book on the table" [topographic]) in contrast to nontopographic sentences (e.g., "The brother is older than his sister"). The left inferior parietal lobule is known to be activated in biological action recognition and in processing the precise configuration and location of hands in space to represent objects, agents, and actions. It has also been shown in other studies to be involved in hand movement imagery when contrasted with actual hand movement (Gerardin et al. 2000) and in imagery of hand rotation (Kosslyn et al. 1998). It is not activated in speech comprehension.

Emmorey et al. (2002) found similar areas of activation in a PET study investigating classifier predicates in ASL (American Sign Language). Deaf signers viewed drawings depicting spatial relations between two objects and were asked either to produce a construction using classifiers, for example, CURVED-OBJECT (the classifier for CUP) signed above FLAT-OBJECT (the classifier for TABLE) or to produce a sentence using an ASL preposition (CUP ON TABLE). In this study the same parietal cortical region was activated as in MacSweeney et al., but analogous right-sided parietal activation was observed as well. Task differences are likely to have driven the different activation patterns in these two studies, because in Emmorey et al. participants had to create sentences in response to images of objects in spatial relations, whereas in MacSweeney et al., participants had only to detect semantically anomalous sentences.

Both studies indicate that some aspects of sign language processing require the contribution of cortical regions not associated with spoken language comprehension. Importantly, in MacSweeney et al., no differential activation in these regions was observed when hearing people heard and saw English translations of topographic BSL sentences. Because the visual medium affords the identification of objects and their spatial locations as a function of their forms and locations on the retina and sensory cortex, it is not surprising that cortical systems specialised for such mappings are utilised when sign languages capture these relationships. The absence of such features in spoken language processing suggests that loss of the parietal link relates to the development of *speech,* rather than to the development of *language,* and thus provides indirect support for Hurford's proposal.

Author's Response

Ventral/dorsal, predicate/argument: The transformation from perception to meaning

James R. Hurford

Language Evolution and Computation Research Unit, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh EH8 9LL, United Kingdom. **jim@ling.ed.ac.uk**

Abstract: It is necessary to distinguish among representations caused directly by perception, representations of past perceptions in long-term memory, the representations underlying linguistic utterances, and the surface phonological and grammatical structures of sentences. The target article dealt essentially with predicate-argument structure at the first of these levels of representation. Discussion of the commentaries mainly involves distinguishing among various applications of the term "predicate"; clarifying the assumed relationship between classical FOPL and language; clarifying the status of unique individuals as conceived by humans; and addressing the issues of motion-perception, binding between object-percepts and predicate-percepts, and target-driven versus stimulus-driven attention.

R1. Introduction: The central claim

The central claim in the target article was that there is a correlation between a fundamental characteristic of any serious formal logical scheme for representing thought and a feature of neural architecture in higher mammals. Every logical scheme has at its heart an asymmetry between two types of terms, usually called *predicates* and *arguments.* Predicates and arguments are essentially different in two ways, namely, in their semantics (how they relate to the world), and in their syntax (how they relate to each other in the formal scheme). In basic logic, arguments denote individual entities, whereas (1-place) predicates denote classes or properties; and the syntax of logic puts predicates outside of (but tied to) the brackets that enclose arguments (though of course other, similarly asymmetric notations are conceivable). This asymmetry, I argued in the target article, finds a parallel in the separation between the ventral and dorsal streams identified in the visual perception systems of higher mammals (and to a lesser extent in their auditory systems).

Note that the above summary of the central claim makes no mention of language. Clearly, non-human animals do not think in language, and much human thought is also not in a public human language, such as English or Chinese. A recurrent theme in many commentaries (**Anderson & Oates**, **Politzer**, **Jones**) was the association of *predicate* with something essentially linguistic. The predicates discussed in the target article are *not* linguistic, they are prelinguistic. Some commentators (**Dessalles & Ghadakpour**, Politzer) usefully distinguished between two senses of *predicate,* a linguistic sense and a non-linguistic sense; these commentaries are thus helpful in clarifying what the target article was (not) about. Jones also noted the different senses of *predicate,* and constructively suggested ways to bridge the gap between theories designed to account for the different phenomena that have been labeled "predicate." Other commentators made no such distinction, and

http://dx.doi.org/10.1017/S0140525X03380070

wrongly assumed that the "predicates" discussed in the target article are identified or closely associated with linguistic categories.

Language is used to express thoughts. As formal logical schemes were developed, largely with human languages as a model, there clearly is a correlation between such schemes and human languages. This makes the over-readiness to identify predicates as linguistic entities understandable. All that the central claim of the target article requires is that it be possible to spell out the nature of the mapping between representations in a logical scheme with the predicate-argument asymmetry at its heart and the grammatical structure evident in languages. The relationship between the surface structure of languages and the representation of thought is so complex and indirect that it will often mask any direct correlation between the surface linguistic structure and the dorsal/ventral separation.

The target article concentrated exclusively on the brain activity involved in perception. It said nothing about how perceived events are stored in memory, except to imply that the predicate-argument asymmetry is presumably not lost, though perhaps somewhat transformed, in this process. When people speak, they mainly express propositions dredged up from memory or generated creatively by recombining elements from different memories. Only very rarely does a human utterance describe what the speaker is perceiving at the very moment of the utterance. Several commentaries implicitly discuss the nature of the representations at the beginning of the utterance production process or at the end of the utterance interpretation process. Such representations are variously called in the literature "semantic representations," "intentional/conceptual representations," "Logical Form," and so forth. Without exception, all proposals for such representations make use of predicate-argument structure. The ubiquity of predicate-argument structure in such representations calls for an explanation, and I attempted to provide one.

R2. Bridging gaps

The relationships between perceptual and long-term memorial representations, and between the latter and linguistic structure, are shown schematically in Figure R1, which will make a useful frame of reference for this reply to commentaries.

Occam's razor dictates that we should try to find theories that minimize the gaps between the boxes in Figure R1. I doubt whether any of the boxes can be eliminated – each seems necessary on independent grounds. Commentaries by **Cowie** and **Jones**, in their very different ways, respectively philosophical and psychological, address the task of bridging the gaps between the boxes in Figure R1.

Cowie asks, "What are the processes of derivation, abstraction, inference, binding, (what to call them?) by which these initial [*PREDICATE(x)*] representations are manipulated or transformed into the kinds of thoughts expressed by our propositional attitude ascriptions?" Dead right – but (speaking of the dead) I do not think that this empiricist research program has been so dead since the 1950s as she supposes. "Empiricism died a death, and rationalism – the idea that there's stuff in our minds that *doesn't* come from the senses (I guess it must be innate!) – emerged as the dominant theoretical orientation throughout the cognitive sciences." This is too dramatic: a winner-take-all Empiricism versus Rationalism scenario is as implausible as the simplistic Nature versus Nurture "debate." In my proposal, the actual distinction between (perceptual) predicates and arguments *is* innate, and doesn't come from the senses. The neuroscientists' discovery of the dorsal/ventral separation simultaneously gives the empiricist philosophical program of relating sense-data to higher-level representations a boost and a cautionary admonition. The innate mechanisms formalizable as *PREDICATE(x)* are *constitutive* of our sense-data, rather than "coming from the senses."

Jones also well appreciates the nature of the gaps between the boxes in Figure R1. His own psychological research has concentrated on the middle and right-hand boxes, those more involved with language, and he has developed a theory that distinguishes between predicates according to their "ease of predication." He asks, "Are the predicate-argument and ease-of-predication approaches compatible? Having been cantilevered out from opposite sides of the cognitive landscape – respectively, from perception/action and from knowledge/language – it seems reasonable to hope that these two stretches of predicational

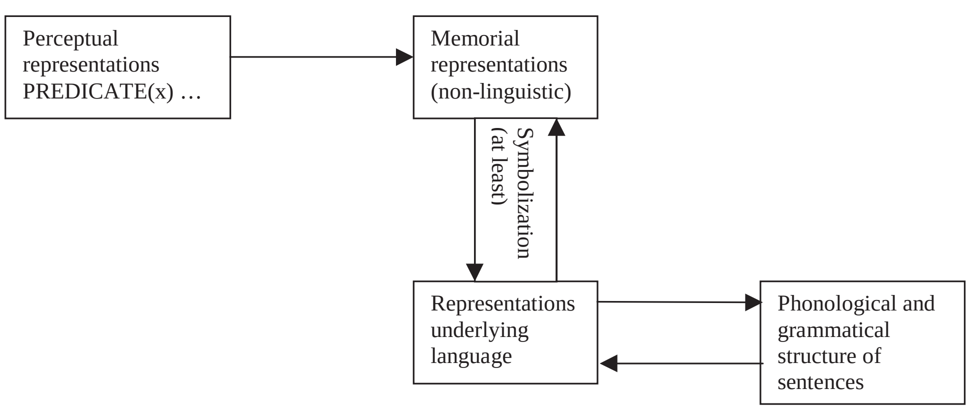

Figure R1. What are the relative "distances" between boxes? What are the relative amounts of processing represented by the arrows? Decreasing the distance between perceptual and memorial representations increases the distance between representations underlying language and surface linguistic structure.

bridge can be made to join up in the middle." I agree, and Jones's analysis of the nature of the problems involved seems acute and accurate.

I have appealed to the idea that predicate-argument structure is somehow transformed, but not beyond recognition, in the translation from perception to memory and thence to underlying linguistic structure. The target article emphasized the logical asymmetry between predicates and individual terms. Several commentators, some apparently quite casually, and some with more deliberation, use formulae that disregard the basic asymmetry, putting predicate terms inside the brackets reserved for argument slots. Examples are "bite(man,dog)" (**Dominey**) and "herd(deer)" (**Lu & Franceschetti**).

Another example, at first sight rather confusing, is **Gillett**'s "in the *PREDICATE(x)* type thought 'that frog is bright orange,' 'that frog' focuses on and tracks an object, and 'bright orange' links a feature of the frog to other stimulus arrays instancing that colour." Here, Gillett has used the predicate term "frog" as an argument of the other predicate "orange." This only coheres in my terms if the proposed representation is not at the perceptual level, where, I hold, representations of the form *PREDICATE(x)* apply. At the perceptual level, I would represent the experience Gillett describes as *FROG(that) & ORANGE(that)* (innocuously using "that" for "x"). What is intriguing and challenging is that representations at a level closer to linguistic form preserve a certain amount of the original asymmetry, but allow the replacement of the argument variables by elements otherwise corresponding to predicates. A theory predicting which predicates can thus "migrate" to the argument slots inevitably raises the issue of the origins of the grammatical noun category, since essentially only nouny predicates have this privilege. You cannot say *That bright orange is frog,* at least not in English. Noun/verb issues recur in these commentaries and are taken up again several times later in this reply.

I sympathize with **Bickerton**'s warning that we do not want to get back to unjustifiably baroque underlying structures for simple sentences. But research has moved on since the heyday of generative semantics, and the target article was concerned with issues other than simply deriving sentences. In particular, the target article focused on perception. As Figure R1 is intended to make clear, there are gaps to be bridged between the representation formed on perceiving an event, the representation of the remembered event in memory, and the structure underlying a sentence used to describe the event. Now, on seeing Floyd breaking some glass, an observer's immediate representation might resemble [*BREAK(e) & FLOYD(x) & AGENT(x) & GLASS(y) & PATIENT(y)*] (always with the caveat that these are non-linguistic predicates). When this event is transferred to long-term memory, it is possible that this representation is transformed into something closer to the linguistic form of *Floyd broke the glass.* It is to be hoped that future research, from both ends of the subject, linguistic and neuroscientific, can shed some light on this question.

Knott also raises the matter of the degree to which the syntactic structures of sentences preserve some characteristics of the raw *PREDICATE(x)* structure. He argues plausibly that such semantically empty grammatical elements as the *there* in *There's a cup on the table* play a similar attention-directing role to the *x* variable in *PREDICATE(x).* Indeed, the etymology in this case is highly suggestive of attention being drawn to a location. This is a topic that deserves a lot more thought, following the idea that when people talk to each other, they are often attempting to recreate experiences in their hearers. Recalling **Gillett**'s example, if, in the absence of the frog, you want to report your *FROG(x) & ORANGE(x)* experience to me, you use the grammatical subject slot as a surrogate for a deictic act of pointing, filling in with a word best calculated to make me think of the right thing. Just why the best term for the purpose is *frog* and not *orange* again raises the noun/verb issue.

Knott's commentary attempts to draw closer parallels than I had envisaged between the structure of perceptual events and grammatical structure. I had relegated the grammatical definite/indefinite distinction to a discourse function, only relevant for communication between people, and playing no part in the solipsistic representation of perceived scenes and events. Knott's idea that there might be some correlate of a definiteness marker in solipsistic representations, indicating something like "entity already known about" is worth following up. The appropriate psychological question would seem to be whether there is any detectable difference between attention directed to a new object for the first time and attention re-directed to an object very recently attended to in the current scene. For example, the kind of attention I might give to a plate, knife, fork, and spoon unexpectedly laid out on a sidewalk might differ from the kind of attention I would pay to these same objects while manipulating them at the dinner table. **Brinck**'s distinction, in his commentary, between objects of attention as either causes, attracting attention, or effects, that is, products of focal attention, may be closely related to Knott's idea. (See below for discussion of Brinck.)

Bridging gaps involves taking advantages of similarities and accounting independently for differences. If two structures have *everything* in common, there is no gap to bridge – they are in fact one and the same structure. **Bridgeman** states that the central claim of the target article "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." Note, first, that "subject and predicate" immediately suggests specifically linguistic structure; the target article acknowledged the differences between perceptual processes and linguistic structure, but claimed that a similar skeletal shape, the asymmetric interaction of elements (processes or terms) of different types, is evident in both. If an archeologist claims that a wooden medieval house was built on the stone foundations of a Roman villa, he may do so on the basis of significant coincidences in the shapes of the separately identified relics. He excavates past the medieval wood to the Roman stone and finds the same basic layout. The archeologist's conclusions do not "miss the mark" because there are obvious differences between the two structures – different materials, at different levels. Metaphorically, the target article was an excavation from the surface structures of languages (dug through pretty quickly), through the logical structures often called *semantic representations,* down to a deep level of mechanisms of visual perception. As with the archeological example, we find a similar layout, in different materials, at the different levels.

I was expecting a more sympathetic hearing from **Bridgeman**, and his commentary seems to have missed some parts of the target article. Bridgeman emphasizes the possibility of contradiction between information in the dorsal and ventral streams. In fact, the target article, in its discussion of blindsight, also acknowledged this; it is one of the characteristics of the perceptual level lacking a clear parallel at the linguistic level. Bridgeman was critical of the use of the terms *what* and *where* for the two pathways. This usage, however inappropriate, is now commonplace, and the target article stuck with it, while explicitly noting that the terms *motor-oriented* and *cognitive* suggested by Bridgeman himself (Bridgeman et al. 1979) are preferable. This terminological point was picked up by **MacNeilage & Davis**. Bridgeman's critique brackets the target article with an earlier paper by Landau and Jackendoff. In fact, the target article explicitly distanced itself from Landau and Jackendoff's (1993) paper, and identified more closely with Bridgeman's own commentary on that paper, emphasizing, in Bridgeman's words, "a level that differentiates linguistic from non-linguistic coding." In this connection, **Piattelli-Palmarini & Harley** also, in mentioning the "chasm between closed- and open-class lexical items," missed the important distance between the target article and Landau and Jackendoff's paper, which it explicitly criticized. Bridgeman also attributes the development of logic to "linguists," and rather generally conflates logic and grammar. He should hear how linguists gossip about logicians – a remark of **Bickerton**'s, quoted below, is an example.

Woll's commentary also raises, in a different way, the gap between perception of events, memory for events, and processing of sentences describing events. Woll wonders why, given my model, the parietal lobes are not involved in linguistic comprehension tasks, especially those that require spatial representational resources. This was briefly addressed in the last paragraph of section 4 of the target article, but some further comment will be helpful, I hope. My claim is about perception – the basic event of having the attention drawn to some object, and then making a judgement about it. The target article said very little about the further processes by which such *PREDICATE(x)* experiences are transferred to memory, perhaps later to be resurrected when talking about such experiences. Clearly the parietal regions need to be busy all the time guiding the attention to objects in the here-and-now. If they were still also active in linguistic production and perception of sentences describing scenes experienced (much) earlier, there could be dysfunctional interference with the second-by-second running of the body parts involved in attention shifts (saccades, hand and head movements).

The following analogy may be useful. Consider a robot programmed to roam the world, taking and storing digital photographs of kinds of things that its programmers have determined. Such a machine would have mechanisms for directing its lens, zooming, focusing, and adjusting exposure for light conditions. Having directed, focused, and adjusted exposure, then "Click!" – the photo is taken and downloaded (say as a JPEG or GIF) to its memory. If the machine is also programmed to search its data bank of photos and to provide descriptive summaries of what they contain, there is no reason to suppose that the lens-directing, zooming, focusing, and exposure-adjusting mechanisms will be involved in this latter task. No analogy is perfect, and probably the brain is not organized in quite the cleanly modular way of this machine: we may not be surprised to find some small residual parietal activity in parsing sentences, especially those about concrete objects in physical space.

R3. Predicates, predicates, . . .

The gaps discussed in the previous section between perception, memory, and representation of language relate to several commentaries which discuss the multiply ambiguous term *predicate.*

Dessalles & Ghadakpour address the difference between "non-linguistic predicate" (their "R-predication") and "linguistic predicate" (their "C-predication"). The target article, although not dwelling on the evolutionary step(s) from one to the other, may have given the impression that a single step, the development of a labeling capacity, is all that is required. Dessalles & Ghadakpour's point about the human capacity for negating and contrasting predicates is important. An animal may be able to represent the concept APPLE, but can it mentally apply a negation operator to get NOT-APPLE? Dessalles & Ghadakpour assume that only humans are capable of this, and we are surely vastly better at it than nonhumans. But there is some evidence that even Alex the parrot can approximate a kind of contrast or negation (Pepperberg 1999). Dessalles & Ghadakpour claim a qualitative difference between R-predication and C-predication, but also concede that object recognition and categorization may be a likely prerequisite of human predication. On the possibility of continuity in the evolution of linguistic predicates from non-linguistic predicates, see **Gillett**'s commentary, discussed briefly later in this response.

Dessalles & Ghadakpour point out that computer simulations such as Batali's (2002) and Kirby's (2000) have predicate-argument structure built in, and this enables the simulated populations of agents to evolve language-like systems. Animals haven't evolved language, Dessalles & Ghadakpour argue, so they can't have predicate-argument structure built in. But an important factor is omitted from their argument here. Batali's and Kirby's simulations did indeed have predicate-argument structure built into the agents (and Batali's *PREDICATE(x)* semantic representations pre-dated my use of them). But the agents also had *symbolic ability,* the capacity to acquire arbitrary labels for their inner concepts. Both this symbolic capacity and the built-in predicate-argument structure are necessary to get such language-like systems to evolve. So my reply to this point of Dessalles & Ghadakpour's is that what is missing in animals is not predicate-argument structure, as they claim, but the rampant symbolizing capacity of humans.

While conforming to the "predicate-as-essentially-linguistic" tendency, **Anderson & Oates** make a different point, which prompts a useful clarification. Their title is "*Prelinguistic agents will form only egocentric representations.*" Predicates, as Anderson & Oates understand the term (their "genuinely objective predicates"), are shared by all members of a community, whereas prelinguistic categorizations are (or may be) idiosyncratic and subjective, and thus "egocentric." I agree. Anderson & Oates argue that the (possible) egocentricity of prelinguistic categorizations, as delivered by the ventral stream, undermines the crucial difference that I rely on between the two information streams. The target article stressed the *deictic* nature of the variable *x* argument in *PREDICATE(x).* It is important to distinguish between *deictic* and *egocentric.* Anderson & Oates are right to say that the agents I envisage are "functionally solipsistic," in that social interaction played no part in my story. But these creatures have a history, and a memory of past events, and a future, and some capacity to plan for it. Animals' mental representations of their pasts and their futures may or may not be functionally solipsistic, but their existence certainly extends beyond the here-and-now. *Deictic* implies "pointing to what is here now." When a creature attends to some object in the here-and-now and categorizes it in such-and-such a way, the category is one that it has carried around in its head ever since it acquired it (or was born with it), and will probably carry it around until its death. The application of a prelinguistic predicate to an object combines ephemeral deictic attention with lasting (if possibly egocentric) mental categories. Prelinguistic predicates may be egocentric, but they are not deictic.

In fact, it is surely likely that prelinguistic creatures of the same species have significantly similar predicates, because of shared adaptive heredity. The work of Luc Steels (1997), to which **Anderson & Oates** refer, starts with agents that have shared constraints on the categorical distinctions they make. The existence of prelinguistic predicates facilitates (is a preadaptation for) the later emergence, aided by language, of what Anderson & Oates call "genuinely objective predicates" and what **Gillett** calls "true concepts." I share the view, clearly expressed by Gillett, that

> convergence in categorization with other competent language users occurs by conversational correction within a colinguistic human group. By noticing this fact, we can, without denying the continuity between human thought and that of higher animals, bring out a point of difference which increases the power of human epistemic activity and in which language plays a central role.

Politzer also distinguishes two senses of *predicate* and *predication,* but his dichotomy differs from that of **Dessalles & Ghadakpour**. Politzer is centrally concerned with predicates as realized in language. " [I]n its modern, logical sense, a predicate (henceforth, 'predicate$_L$') is a function from a singular term to a sentence expressing a proposition about the object to which the singular term refers." This is somewhat close to my sense in the formula *PREDICATE(x).* But note that in writing about sentences, and therefore about *linguistic* predicates, Politzer is implicitly granting the transition from mental prelinguistic predicates to (this type of) linguistic predicates, a transition that others found problematic or at least in need of an account. The definition of predicate$_L$, Politzer comments, is at odds with "Aristotelian sentences" ("A-sentences") which, as the target article noted, do not require the subject and the grammatical predicate ("predicate$_G$") to belong to different kinds of terms. Examples: *The philosopher is a man* and *The man is a philosopher.*

The subjects of modern human sentences, **Politzer** argues, can be interpreted either intentionally or extensionally. As a piece of purely synchronic description, this may be adequate. But Politzer's suggested account of the origin of A-sentences is circular, resting implicitly on an unstated distinction between nouns and adjectives. He writes

> From the inception of categorization, the A-sentence predication could start to develop, taking *generic* individual objects as its subject: "x is a predator" (with temporal anteriority) and "x is yellow" are conflated into "the predator is yellow." Indeed, it would be uneconomical to formulate, for example, *the x is S and P* when the first predication is temporally or cognitively already established; hence the shorter formulation *the S is P.*

Politzer's appeal to "temporal anteriority" is crucial to his argument and quite unjustified. What is to guarantee that the judgement that something is a predator precedes the judgment that something is yellow? One would in fact tend to expect the opposite. Without this ad hoc appeal to anteriority, there is no way to avoid *the yellow is predator.* Politzer seems implicitly to have accorded "anteriority" to predicates corresponding to nouns, but gives no reason why only noun-predicates are "established first." Likewise, the psychological experiment on syllogisms to which Politzer appeals reflects an unwillingness in subjects to concoct sentences in which nouns and adjectives are in the wrong grammatical positions. Clearly, these arguments do not explain how a logically uniform class of predicates came to be mapped onto grammatically contrasting syntactic categories – nouns versus others; the Aristotle problem is still unsolved. Finally, on a conciliatory note, Politzer may be right that certain kinds of predicate have "anteriority," although he himself gives no account of what such anteriority is. Perhaps something like **Jones**'s ease of predication (see above) does in fact distinguish noun-predicates from others. None of this critique of Politzer's commentary affects the central claim of the target article, correlating *PREDICATE(x)* with the ventral/dorsal separation, a claim that Politzer finds plausible.

Anastasio comments on the inappropriateness of discrete, categorical, logical representations, like *PREDICATE(x),* and contrasts formal logic with probability. I agree with Anastasio's main point, that a properly fine-grained account of brain activity has to be in terms of probabilities. Anastasio gives the example of what happens in the ventral stream when someone recognizes an object as satisfying the predicate *APPLE,* emphasizing that the relevant cortical neurons are not two-state elements, but show graded responses to their inputs. This is absolutely right, as is **Arbib**'s similar comment that neural schemata are likelihood distributions, rather than simply returning 1 or 0. As the target article stated, the brain is vastly more complex and subtle than any formal scheme invented by a logician.

But we should not, and **Anastasio** does not, throw out the predicate-argument baby with the discrete, categorical bathwater. Anastasio suggests replacing a predicate-argument formula such as *APPLE(x)* with $P(X=\text{APPLE}|S)$, representing the probability that an object *X,* eliciting sensory input *S,* is an apple. This models the activity of a neuron in the ventral stream. The predicate-argument distinction is still present in Anastasio's formula, in the form of the two terms *APPLE* and *X,* merely hedged around with a probability factor. (I would prefer to replace Anastasio's "equals" sign with something less suggestive of identity and conveying the relation of satisfaction between a predicate and its argument.) And Anastasio has no quarrel with the target article's correlation of predicate and argument with the ventral and dorsal streams. Logic and probability are not incompatible. Probabilistic, or "fuzzy" logics have been developed (see, e.g., Zadeh & Kacprzyk 1992).

R4. Logic

"'Logic' is not so well defined a term, nor logic so tidy or static a discipline, as the popular conception of the logician as a paradigmatically convergent thinker minding his Ps and Qs might lead one to suppose." (Haack 1994, p. 891). A theme in some commentaries is a mild antipathy to logicians and all their works. For example, **Bickerton** writes,

http://dx.doi.org/10.1017/S0140525X03380070

"sentences of a type seldom uttered by non-logicians ('Socrates is a man')", and "what were our remote ancestors most concerned about, getting their FOPL straight or telling one another interesting things?" The target article is in no way a claim about the whole apparatus of first order predicate logic (FOPL), and still less about the whole range of (often incompatible) models of logic found in the literature. The central claim is only about the predicate-argument asymmetry at the heart of logical structure. Commentators have taken no exception to my occasional appeals to the conjunction of predicate-argument formulae, or to the implicit existential quantifier binding the *x* of the *PREDICATE(x)* formula. But I do not claim any primitive status for other features of FOPL, such as logical disjunction or universal quantification. And the target article stated clearly that the individual constants of FOPL are "practically unrealistic, requiring Godlike omniscience." The logician's disciplined insistence on a well-specified ontology mapped explicitly onto elements of the logical notation is exemplary. This complete rigor allows precise evaluation of aspects of the system in psychological terms, prompting us to reject logical individual constants (rigid designators). More generally, such a mapping (or a "model" or a "denotation assignment function") is necessary in any serious theory about how people can tell one another interesting things.

The target article appealed to a broad class of semantic theories such as Montague Grammar, event semantics, and discourse representation theory, which have made plausible attempts to relate quite abstract, and often typically flat, semantic representations to the surface grammar of languages. It is just these semantic theories that **Dominey** characterizes as "developing a theoretical basis for mapping logic to language and the meanings that can be expressed in language." But this misrepresents the goals of these theories. They aim to map the structures found in language onto a type of well-specified external world (a "model"). They attempt to provide an account of what natural language sentences are about. Even Frege, in his way, was trying to do just this. Likewise, **Bickerton** attributes to me "an assumption that language and cognition must be based on logic" and an "insistence on approaching language from a logical point of view." **Carstairs-McCarthy**, very similarly, claims I assume "that there *must* be some stage of linguistic evolution at which (proto)syntax behaved in a fashion that reflected more closely than it does now the way in which predicate-argument structure works in logic" (emphasis added). I do not assume this a priori appropriateness of logic. The work of logicians is to be respected for its rigor, and appreciated when we can discover in logic something that is useful in an account of cognition and language. I plead guilty to cherry-picking what is useful from logic. Hence my *rejection* of individual constants (as defined in logic) in favor of individual variables, which, I claim, can be correlated quite nicely with the information from the dorsal stream. A logical notation is not an a priori Procrustean starting point, with which semanticists wrestle to make language map onto it. This is just why **Cowie** sees the target article as *vindicating* the predicate-argument asymmetry in logic, rather than the other way around. Nowadays the logical notations that semanticists use are theoretical constructs designed with the goal of relating sentences to the world they describe (and in some cases to further functions of sentences as well). They are driven by language – not by the structures of isolated sentences, but by the complex semantic interrelationships between whole classes of diversely structured sentences.

R5. Semantic representations

The tenor of my proposal was to assume that in the transformation from perception to memory, the basic ontology of the predicate and argument terms is preserved, respecting the distinction between individual variables and constant or "universal" properties. This assumption then implies "flat" conceptual representations of propositions like *WRIGGLE(x) & BROWN(x) & WORM(x)*, which could get expressed in English as *A brown worm wriggled* or *There was something brown and wriggling; it was a worm,* and so on. I am interested in making a case for such flat representations, but they were not a central plank of the target article. It might turn out to be more correct that the representations that we parse into, and that we start from when speaking, are closer to the surface grammar of languages, so that the original perceptual/ontological distinctions are lost, as they are in a representation such as *wriggle(worm),* where the predicate *WORM* has been translated into a term capable of occupying an "argument" slot. But notice that any narrowing of the gap between semantic representations and the surface grammar of languages correspondingly will tend to widen the gap between semantic representations and the raw experiences of perception on which (I claim) they are based.

Dominey's proposal is to ditch the work of formal semanticists because of the quantity of effort expended. He rejects a flat structure *bite(e), man(x), dog(y), agent(x), patient(y)* as a representation of *A man bites a dog,* on the grounds that it is arbitrary, unordered, and less informative than *bite(man,dog).* The flat structure is not "arbitrary"; on the contrary, it is designed to account economically for a whole range of facts about the sentence in question, facts such as the following. The sentence is a paraphrase of others such as *A dog is bitten by a man, It is a dog that the man bites, What bites the dog is a man.* Of the situation described by the sentence, the question can be asked, "Who (or what) acted deliberately?" eliciting the Agent information separately, and "Who (or what) was affected by the act?" eliciting the Patient information separately. The semantic structure of a sentence is not a matter that can be determined solely by looking at it in isolation. A proposed representation such as *bite(man, dog)* needs to be backed up by thoroughly reasoned comparison with the alternatives, taking into account a worked-out theory of what in the world the terms denote. How, for example, will Dominey treat the term *man* in dealing with a sentence such as *Peter is a man*? Will *man* be assigned the same denotation for both sentences, and if so what is it?

Obviously, at one end of the theoretical derivation of an individual sentence there has to be a structure that closely resembles the sentence itself. But once we undertake to relate whole sets of sentences to each other, to the eventualities they describe, and to the psychological mechanisms that process both the sentences and the eventualities, things get more complicated. In terms of the box diagram in Figure R1, **Dominey**'s proposed representations are closer to the right-hand linguistic end. Dominey appears to accept the case for relating neural architecture to *PREDI-*

CATE(x) structure, but rejects the possibility of mapping *PREDICATE(x)* structures onto language. But this leaves a gap unbridged.

An analogy might help us out of our difficulty here. Getting (digital) computers to work involves converting everything down to a binary code consisting of nothing but 1s and 0s. Fortunately for users, intermediate levels of representation, such as machine code, assembly language, and high-level languages, keep us from toiling with 1s and 0s. Anything can be coded in 1s and 0s, and reliably retrieved. With computers, the effort is necessary, because the basic electronics deals best with 1s and 0s. A theory of the input-output mappings of a particular programmed computer might not need to get down to the 1s and 0s, but if the mappings were at all complicated, some quite abstract underlying representations, no doubt resembling the computers' program(s), would be required. The *PREDICATE(x)* formula is somewhere in the middle between the firings of individual neurons (analogous to 1s and 0s) and our representations of the inputs and outputs that humans map onto each other, namely sentences and perceptions of objects and events. Just as anything, of whatever dimensionality, can be coded into a linear one-dimensional representation, even with the minimal {1, 0} alphabet, anything can be coded into flat structures of conjoined *PREDICATE(x)* elements. The question of whether this is the most elegant solution is open, but cannot be easily dismissed.

Dominey writes that "Prelingual infants appear to represent collisions in terms of the properties of the 'collider' and their influence on the 'collidee.' This supports (but does not prove) the hypothesis that contact is represented by a 2(or greater)-place predicate." Just the contrary, surely. The target article made a brief case that 2-place predications can always be reduced to conjunctions of 1-place predications assigning properties differentially to the participants. And here are the prelingual infants conforming to that conjecture.

The question of whether (perception of) all eventualities can be represented by conjunctions of 1-place predications, or whether some 2-place predications are necessary is also raised by **Werning** and **Arbib**, but in a different way from **Dominey**. Like me, and unlike Dominey, both Werning and Arbib use flat conjunctions of predications with individual variables as their arguments, as in *bite(e), man(x), dog(y), agent(x), patient(y)*. Arbib writes "I don't think this works. We need to replace *agent x* by *agent(x, e)* to indicate in which event *x* plays the stipulated role." For Werning, too, the predicates representing participant roles, such as Agent and Patient, are 2-place, for the same reason. This is common in neo-Davidsonian event semantics. I am interested in pursuing the possibility that such 2-place predications can be eliminated by bracketing together all the predications relating to a single event, and restricting the application of such predicates as *agent* and *patient* to the local environments delimited by the brackets. The box notations of Discourse Representation Theory (Kamp 1981; Kamp & Reyle 1993), or something like them, could perhaps be used for this role-scope-delimiting purpose.

Lu & Franceschetti, following Talmy (2000), point to the need for the components Figure, Motion, Path, and Ground in semantic representations. This seems right. The target article discussed some similar examples to theirs, in section 1.2, but without using Talmy's terminology. Lu & Franceschetti's discussion is quite informal, but there is no obvious reason why, in a more formal representation, these semantic components could not be expressed as 1-place predicates applying to the various objects in the situations and events described.

Werning clearly accepts the standard event-semantic arguments for the kind of flat-style semantic representations, with variables for arguments, that I assume. Thus we tend in the same direction in how far we distance semantic representations from the surface grammar of languages. But Werning makes one concession that I do not make to language-like structure. His semantic representations distinguish between two kinds of variable occupying the argument slots, namely, object variables and event variables, and these are used to explain the universal linguistic distinction between nouns and verbs. Werning, like me, is committed to finding neural correlates for semantic representations. He, unlike me, associates predicates over event variables, giving rise to linguistic verbs, with the ventral stream, and predicates over object variables, giving rise to linguistic nouns, with the dorsal stream. I will discuss the important issue of brain areas involved in detection of motion in another section. But note here that Werning's account loses the opportunity to explain the predicate-argument asymmetry. If the *SLUMP(e)* of his example, paraphraseable as *There was a slumping event,* is completely hosted by the dorsal stream, what accounts for the predicate-argument structure that Werning assigns to it? Likewise, if *RED(x)* is entirely hosted by the ventral stream, what accounts for the predicate-argument structure Werning assumes it to have? I am not claiming that my correlation of the ventral/dorsal separation must be the only possible explanation for the predicate-argument asymmetry, but it is at least one proposed explanation. Werning's account leaves unanswered the question of how, in each stream separately but in parallel, the same basic predicate-argument asymmetry arises.

R6. Individuals and linguists

Interestingly, of the three commentaries by linguists, two (by **Bickerton** and **Carstairs-McCarthy**) concentrated on issues to do with individuals and proper names. Their comments in some ways echo those of **Dominey**, **Politzer**, and **Gillett**, discussed earlier. It is useful to clarify what the separate issues here are.

Can humans reliably reidentify unique individuals? As **Carstairs-McCarthy**'s example of his possible confusion of Jim Hurford with Tim Hurford shows, we cannot. We are in no better position than the tern chicks, as both human and non-human animals, and we are at the mercy of our senses for reidentifying particulars.

Do humans have concepts of unique individuals? I agree with **Carstairs-McCarthy** that we do. He is surely right in stating that a Matsigenka who happens to have two relatives glossed as "patrilineal same-sex cousin" will have concepts of these relatives as separate unique individuals. However, I submit that evidence for the human concept of a unique individual rests only on our human command of the meaning of *unique* and its translations in other languages, and our grammatical intuitions. Since we understand the meaning of the word *unique,* we can hardly deny that we have the corresponding concept. The only other evidence that humans have concepts of unique individuals comes from our grammatical intuitions. We intuit such facts as that *Mary* and *herself* are "co-referential" in *Mary admires herself.*

There is a curious tension between the answers to the two questions posed above, parallel with a tension between ideal linguistic competence and actual linguistic performance. In a previous paper (Hurford 1999) I suggested a process by which certain perceived individuals, by virtue of constant presentation of complex combinations of properties salient and important for us, become "cognized individuals," attributed with uniqueness in our belief system. In the terms of Figure R1, certain perceptions get transferred to memory with such frequency and significance for us, that the *whichness* of the objects concerned is preserved. In the words of that paper (reidentifiable) "individuals are abstractions." Further, by our constant use of language that presupposes and bolsters our beliefs in unique individuals, we may construct and reinforce mental representations containing a certain class of term that, we believe, is used to reidentify unique individuals.

Do non-human animals have concepts of unique individuals? The concept of uniqueness rests on a command of negation. A full command of uniqueness cannot exist without a command of negation and contrast. **Dessalles & Ghadakpour** argue that non-human animals lack a command of negation and contrast, and I agree that such concepts are likely to be less developed in non-human animals. There is no aspect of animals' behavior forcing us to attribute concepts of unique individuals to them, as opposed to a more parsimonious attribution of sufficiently specific conjunctions of properties. The signing chimpanzees that **Carstairs-McCarthy** mentions use signs glossed as *Roger* and *Washoe.* But that does not immediately tell us that these animals had a mental conception of Roger, for example, as unique in a way that the denotata of *tree, cage,* and *house* did not. (I keep quiet about *Washoe,* as there is no space to get into the thorny issue of animals' conceptions of themselves as unique or otherwise.) The question is the extent to which repeated exposure to and familiarity with certain things can give rise, in animals, to a special class of concepts of cognized individuals, each member of which is credited with uniqueness. I surmise that such concepts came late in human evolution – maybe before something approaching full human language (and gossip), maybe not. We certainly did not need them to stand for the forerunners of arguments in predicate-argument structures, because the availability of *perceived* individuals, as opposed to cognized individuals, was there from long before humans evolved. **Bickerton**'s assertion that "they [animals] have a clear concept of a specific individual" is not backed by evidence. We may fondly believe that our pet cats treat us as special individuals until we see them sidle purring around the next houseguest who feeds them.

Do languages have proper names? Yes, obviously, most do, and some, like Matsigenka, do not. How do we decide whether a language has proper names? The standard linguistic method for answering such a question appeals to the distribution of expressions in well-formed sentences. **Bickerton**'s hypothetical language has two sentences: *Last night the wind knocked the hut over* and *Knocked the hut over seduced your wife last night.* Bickerton claims that in the second sentence, but not the first, *knocked the hut over* is a proper name. This is not, of course, enough data to go on, but we get the idea. I suppose that on similar grounds, he might conclude that *the teacher* is a proper name in *The teacher should know better* whereas the same expression is not a proper name in *Derek is the teacher.* But of course, in both sentences, *the teacher* belongs to the same syntactic category, NP, which is not equivalent to proper name. (The copula *is* here is not crucial, as many languages, e.g., Arabic, Russian, express predication without such a connecting particle.) Bickerton might disagree: "What determines whether something is a proper name is not its internal structure but how it is used." This might reflect the assumption that "proper name" is a semantic, rather than syntactic, category, so in this case "proper name" means "expression used to refer to an individual." Even in languages without proper names, it is (unsurprisingly) possible to refer to individuals. The question is whether the precursors of language had available a special category dedicated to this purpose.

(There are several odd misconceptions in **Carstairs-McCarthy**'s commentary. His abstract states that I claim that the priority of empty variables in predicate-argument structure "had an effect on visual perception." In fact, of course, I claimed just the opposite. He also implies that "proper names are complex to handle in first-order predicate logic." Semanticists attempt to account for the mapping between a large body of natural language (typically English) sentences and models of a world that they describe; what they find hard to handle is dealing with linguistic proper names as straightforward equivalents of individual constants in FOPL.

The other commenting linguists were **Piattelli-Palmarini & Harley**. They also discussed individuals, but in connection with a different question. They wonder why the clauses found in natural languages are typically restricted to a maximum of three or four principal participants. The target article (sect. 4) attributed this directly to nonlinguistic limitations of short-term memory, as surveyed by Cowan (2001). This explanation is not addressed by Piattelli-Palmarini & Harley, who prefer an explanation in terms of language-internal constraints. But such an appeal to language-internal constraints leaves unanswered the question of where such constraints came from. I have a preference for reductionist explanations; Piattelli-Palmarini & Harley prefer a "translationist" project. I would like to know whether research programs of this "translationist" stripe could ever explain phenomena of emergence or evolution.

R7. Events, motion, and the dorsal and ventral streams

For several commentators, the structure of perceived events and the perception of motion are a locus of problems for my approach. **Lu & Franceschetti** discuss the perception of events without specific mention of the neural processing streams involved. They appeal to psychological studies that analyze the stream of motion as constructed of basic building blocks that are temporal units in which Figure, Motion, Path, and Ground are constant. A change in any of these features constitutes a new event. As suggested briefly above, it may be possible to represent each such unitary temporal building block as a conjunction of 1-place predications involving Figure, Motion, Path, and Ground as properties of the participants. Also following an earlier suggestion, a box as in the notation of Discourse Representation Theory (DRT) could bracket each such conjunction. Changes of state, that is, events, could possibly be represented in the DRT fashion by temporal indexing of separate boxes. Such notation-juggling does not, however, engage 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 perhaps just using, or even being used by) some configuration in its brain. Perhaps **Piattelli-Palmarini & Harley**'s point is that whatever is accessed internally in this case is not a "description." I do not claim, of course, that if we open up a brain we will find representations somehow written down in the same kind of public symbols that we humans use to talk about things (any more than we will find a little homunculus looking at a screen). Piattelli-Palmarini & Harley write "Neural states or processes as such have no semantics. They co-vary nomologically and causally with events in the world." My view is that the causal co-variance of neural states or processes with events in the world is a necessary but not sufficient basis of semantics. Smoke is causally co-variant with fire. To a first approximation, semantics, as the term is conventionally used, is restricted to the domain of conventional or non-natural meaning, in Grice's terms.

I agree with **Arbib** that animals' representations are modulated in complex ways by whatever else is happening in the brain and in the world outside; so such representations are indeed "likelihood distributions over a multi-dimensional parameter space." Since Wittgenstein, most semanticists have believed that the meanings of words are also likelihood distributions over a multi-dimensional parameter space, so the use of an expression like "non-linguistic predicate" should not be too objectionable.

R12. Where next?

I thank *BBS* and the commentators for the opportunity to air these ideas. Valid points have been made relating to the central claim of the target article, but I believe there is still insight to be gained from developing it and exploring its ramifications. Neuroscientific exploration, psychological experimentation, and formal semantic work should proceed in parallel on a range of topics, including the basis of the noun/verb distinction, the nature of events and motion, the relationship between objects and events, the relation between perception and memory, the relation between non-linguistic memorial representations and linguistic structure, and different kinds of attention. A truly unified account of linguistic behavior will require the active engagement of scholars from very diverse traditions. Linguists and logicians will need to get more familiar with the neuroscience literature, and neuroscientists and psychologists will need to develop a better understanding of the methods and concerns of those working in more formal traditions. *Behavioral and Brain Sciences* is an excellent journal in promoting just this kind of interdisciplinary exchange.

References

Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively.

Aglioti, S., Goodale, M. A. & DeSouza, J. F. X. (1995) Size-contrast illusions deceive the eye but not the hand. *Current Biology* 5:679–85. [aJRH]

Agre, P. E. & Chapman, D. (1987) Pengi: An implementation of a theory of activity. *Proceedings of the AAAI* 87:268–72. [aJRH]

Aguirre, G. K. & D'Esposito, M. (1997) Environmental knowledge is subserved by separable dorsal/ventral neural areas. *The Journal of Neuroscience* 17(7):2512–18. [aJRH]

Aitchison, J. (1996) *The seeds of speech.* Cambridge University Press. [BI]

Anastasio, T. J., Patton, P. E. & Belkacem-Boussaid, K. (2000) Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Computation* 12:997–1019. [TJA]

Andersen, R. A., Bracewell, R. M., Barash, S., Gnadt, J. W. & Fogassi, L. (1990) Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7A of macaque. *Journal of Neuroscience* 10:1176–96. [aJRH]

Andersen, R. A., Essick, G. K. & Siegel, R. M. (1985) Encoding of spatial location by posterior parietal neurons. *Science* 230:456–58. [aJRH]

Anderson, S. R. & Keenan, E. L. (1985) Deixis. In: *Language typology and syntactic description, vol. III: Grammatical categories and the lexicon,* ed. T. Shopen. Cambridge University Press. [aJRH]

Antell, S. E. & Keating, J. P. (1983) Perception of numerical invariance in neonates. *Child Development* 54:695–701. [aJRH]

Arbib, M. A. (1972) *The metaphorical brain: An introduction to cybernetics as artificial intelligence and brain theory.* Wiley-Interscience. [MAA]

(1981) Perceptual structures and distributed motor control. In: *Handbook of physiology, section 2: The nervous system, vol. 2, motor control, part 1,* ed. V. B. Brooks. American Physiological Society. [MAA]

(1987a) Levels of modeling of visually guided behavior (with peer commentary and author's response). *Behavioral and Brain Sciences* 10:407–65. [MAA]

(1987b) Modularity and interaction of brain regions underlying visuomotor coordination. In: *Modularity in knowledge representation and natural language understanding,* ed. J. L. Garfield. MIT Press. [aJRH]

Arbib, M. A., Érdi, P. & Szentágothai, J. (1998) *Neural organization: Structure, function, and dynamics.* MIT Press. [MAA]

Armstrong, D. M. (1978) *Nominalism and realism: Universals and scientific realism, vol. I.* Cambridge University Press. [aJRH]

Ballard, D. H., Hayhoe, M. M. & Pelz, J. B. (1995) Memory representations in natural tasks. *Journal of Cognitive Neuroscience* 7(1):66–80. [aJRH, AK]

Ballard, D. H., Hayhoe, M. M., Pook, P. K. & Rao, R. P. N. (1997) Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20(4):723–67. [IB, aJRH, AK]

Barlow, H. B. (1969) Pattern recognition and the responses of sensory neurons. *Annals of the New York Academy of Sciences* 156:872–81. [TJA]

(1972) Single units and sensation: A neuron doctrine for perceptual psychology? *Perception* 1:371–94. [TJA]

Barrett, M. (1995) Early lexical development. In: *The handbook of child language,* ed. P. Fletcher & B. MacWhinney. Blackwell. [BI]

Batali, J. (2002) The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In: *Linguistic evolution through language acquisition: Formal and computational models,* ed. T. Briscoe. Cambridge University Press. [MLA, J-LD, aJRH]

Baylis, G. C. (1994) Visual attention and objects: Two-object costs with equal convexity. *Journal of Experimental Psychology: Human Perception and Performance* 20(1): 208–12. [aJRH]

Baylis, G. C. & Driver, J. (1993) Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance* 19(3):451–70. [aJRH]

Bear, M. F., Connors, B. W. & Paradiso, M. A. (2001) *Neuroscience: Exploring the brain.* Lippincott. [SL]

Beintema, J. A. & Lappe, M. (2002) Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences USA* 99(8):5661–63. [rJRH]

Belin, P. & Zatorre, R. J. (2000) 'What', 'where' and 'how' in auditory cortex. *Nature Neuroscience* 3(10):965–66. [aJRH]

Bellugi, U, Sabo, H. & Vaid, J. (1988) Spatial deficits in children with Williams syndrome. In: *Spatial cognition: Brain bases and development,* ed. J. Stiles-Davis, M. Kritchevsky & U. Bellugi. Erlbaum. [aJRH]

Bennett, J. (1976) *Linguistic behavior.* Cambridge University Press. [aJRH]

Berkeley, G. (1710/1962) *The principles of human knowledge.* Open Court. [FC]

Bickerton, D. (1998) Catastrophic evolution: The case for a single step from protolanguage to full human language. In: *Approaches to the evolution of language: Social and cognitive bases,* ed. J. R. Hurford, M. Studdert-Kennedy & C. Knight. Cambridge University Press. [aJRH]

(1990) *Language and species.* University of Chicago Press. [aJRH]

Bishop, C. M. (1995) *Neural networks for pattern recognition.* Clarendon Press. [TJA]

Blaser, E., Pylyshyn, Z. W. & Holcombe, A. O. (2000) Tracking an object through feature space. *Nature* 408:196–99. [aJRH]

Bloom, L. (1973) *One word at a time.* Morton. [BI]

Bloom, P. (2000) *How children learn the meanings of words: Learning, development and conceptual change.* MIT Press. [aJRH]

Borer, H. (in press) *Structuring sense.* University of Southern California. [MP-P]

Bridgeman, B. (1992) On the origin of consciousness and language. *Psycoloquy* (refereed electronic journal) March 17, 1992. http://psycprints.ecs.soton.ac.uk/archive/00000239/ [BB]

(1993) Spatial and cognitive vision differentiate at low levels, but not in language. *Behavioral and Brain Sciences* 16(2):240. [aJRH]

(1999) Vertical modularity in the visual system. In: *Stratification in cognition and consciousness,* ed. B. M. Velichkovsky & B. Challis. John Benjamins. [BB]
Bridgeman, B., Gemmer, A., Forsman, T. & Huemer, V. (2000) Properties of the sensorimotor branch of the visual system, *Vision Research* 40:3539–52. [BB]
Bridgeman, B., Lewis, S., Heit, G. & Nagle, M. (1979) Relationship between cognitive and motor-oriented systems of visual position perception. *Journal of Experimental Psychology: Human Perception and Performance* 5:692–700. [aJRH, PFM]
Bridgeman, B., van der Heijden, A. H. C. & Velichkovsky, B. M. (1994) A theory of visual stability across saccadic eye movements. *Behavioural and Brain Sciences* 7(2):247–92. [aJRH]
Brinck, I. (2001) Attention and the evolution of intentional communication. *Pragmatics and Cognition* 9(2):255–72. [IB]
Brooks, R. (1991) Intelligence without representation. *Artificial Intelligence* 47:139–59. [IB]
Brotchie, P. R., Anderson, R. A., Snyder, L. H. & Goodman, S. J. (1995) Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature* 375:232–35. [aJRH]
Bryant, D. J. (1993) Frames of reference in the spatial representation system. *Behavioral and Brain Sciences* 16(2):241–42. [aJRH]
Calvin, W. H. & Bickerton, D. (2000) *Lingua ex machina: Reconciling Darwin and Chomsky with the human brain.* MIT Press. [aJRH]
Campbell, J. (1994) *Past, space, and self.* MIT Press. [GG]
Campbell, R. & Woll, B. (2003) Space is special in sign. *Trends in Cognitive Sciences* 7:1:5–7. [BW]
Carnap. R. (1929/1967) *The logical structure of the world and pseudoproblems in philosophy.* University of California Press. [FC]
(1939/1949) *The logical syntax of language.* Routledge. [FC]
(1942) *Introduction to semantics.* MIT Press. [aJRH]
Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F. & Thulborn, K. R. (1999) Timecourse of fMRI activation in language and spatial networks during sentence comprehension. *Neuroimage* 10:216–24. [BW]
Carrozzo, M., McIntyre, J., Zago, M. & Lacquaniti, F. (1999) Viewer-centered and body-centered frames of reference in direct visuomotor transformations. *Experimental Brain Research* 129(2):201–10. [aJRH]
Carstairs-McCarthy, A. (1999) *The origins of complex language: An inquiry into the evolutionary beginnings of sentences, syllables and truth.* Oxford University Press. [AC-M]
Cassirer, E. (1955) *The philosophy of symbolic forms, vol. 1: Language,* trans. R. Manheim. Yale University Press. [BI]
Chen, J., Myerson, J., Hala, S. & Simon, A. (2000) Behavioral evidence for brain-based ability factors in visuospatial information processing. *Neuropsychologia* 38:380–87. [aJRH]
Cheney, D. L. & Seyfarth, R. M. (1990) *How monkeys see the world: Inside the mind of another species.* University of Chicago Press. [AC-M]
Chomsky, N. (1988) Language and interpretation: Philosophical reflections and empirical inquiry. [MP-P]
Clarke, S., Bellmann, A., Meuli, R. A., Assal, G. & Steck, A. J. (2000) Auditory agnosia and auditory spatial deficits following left hemispheric lesions: Evidence for distinct processing pathways. *Neuropsychologia* 38:797–807. [aJRH]
Copi, I. (1958) Artificial languages. In: *Language, thought and culture,* ed. P. Henle. University of Michigan Press. [aJRH]
Cowan, N. (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24(1):87–114. [aJRH]
Croft, W. (2000) Parts of speech as language universals and as language-particular categories. In: *Approaches to the typology of word classes,* ed. P. M. Vogel & B. Comrie. Mouton de Gruyter. [MW]
Cruse, A. (1973) Some thoughts on agentivity. *Journal of Linguistics* 9:11–23. [aJRH]
Culham, J. C. & Kanwisher, N. G. (2001) Neuroimaging of cognitive functions in human parietal cortex. *Current Opinion Neurobiology* 11:157–63. [BW]
Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D. & Damasio, A. (1996) A neural basis for lexical retrieval. *Nature* 380:499–505. [PFM]
Damasio, H., Grabowski, T. J., Tranel, D., Ponto, L. L. B., Hichwa, R. D. & Damasio, A. R. (2001) Neural correlates of naming actions and naming spatial relations. *NeuroImage* 13:1053–64. [BW]
Darwin, C. (l871) *The descent of man.* Great Books, Encyclopedia Brittanica. [PFM]
Davidson, D. (1980) *Essays on actions and events.* Clarendon Press. [aJRH]
DeGelder, B., Vroomen, J., Pourtois, G. & Weiskrantz, L. (1999) Non-conscious recognition of affect in the absence of striate cortex. *NeuroReport* 10(18)3759–63. [aJRH]
(2000) Affective blindsight: Are we blindly led by emotions? *Trends in Cognitive Sciences* 4(4):126–27. [aJRH]
Dehaene, S. (1997) *The number sense.* Oxford University Press. [aJRH]
De Mornay Davies, P. & Funnell, E. (2000) Semantic representation and ease of predication. *Brain and Language* 73:92–119. [GVJ]
Dessalles, J.-L. (1996) *L'ordinateur genetique.* Hermes. [J-LD]
(2000) *Aux origines du langage: Une histoire naturelle de la parole.* Hermes. [J-LD]
Didday, R. L. & Arbib, M. A. (1971) The organization of action-oriented memory for a perceiving system I: The basic model. *Journal of Cybernetics* 1:3–18. [MAA]
Dixon, R. M. W. (1980) *The languages of Australia.* Cambridge University Press. [aJRH]
(1982) *Where have all the adjectives gone?: And other essays in semantics and syntax.* Mouton. [aJRH]
Dominey, P. F. (2000) Conceptual grounding in simulation studies of language acquisition. *Evolution of Communication* 4(1):57–85. [PFD]
(2002) Language learning for human-machine interaction via mapping of grammatical structure to visual scene structure. Institut des Sciences Cognitives (ISC) Working Paper 2002–2, http://www.isc.cnrs.fr/wp/DomineyISC.PDF. [PFD]
(2003) Learning grammatical constructions in a miniature language from narrated video events. In: *Proceedings of the 25th Annual Meeting of the Cognitive Science Society.* (in press). [PFD]
Dominey, P. F., Hoen, M., Blanc, J. M. & Lelekov-Boissard, T. (2003) Neurological basis of language and sequential cognition: Evidence from simulation, aphasia and ERP studies. *Brain and Language* [PFD]
Douglas, K. (2001) Playing fair. *New Scientist* 2281:38–42. [aJRH]
Dowty, D. (1991) Thematic proto-roles and argument selection. *Language* 67(3):547–619. [aJRH]
Driver, J. & Baylis, G. C. (1998) Attention and visual object segmentation. In: *The attentive brain,* ed. R. Parasuraman. MIT Press. [IB]
Duda, R. O., Hart, P. E. & Stork, D. G. (2001) *Pattern classification.* Wiley. [TJA]
Duhamel, J.-R., Colby, C., & Goldberg, M. (1992) The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255:90–92. [aJRH]
Dunbar, R. I. M. (1993) Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences* 16:681–735. [DB]
Duncan, J. (1984) Selective attention and the organization of visual information. *Journal of Experimental Psychology: General* 113(4):501–17. [aJRH]
Egly, R., Driver, J. & Rafal, R. D. (1994) Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General* 123(2):161–77. [aJRH]
Eilan, N. (1998) Perceptual intentionality, attention and consciousness. In: *Contemporary issues in the philosophy of mind,* ed. A. O'Hear. Cambridge University Press. [IB]
Eimer, M., Nattkemper, D., Schröger, E. & Prinz, W. (1996) Involuntary attention. In: *Handbook of perception and action, vol. 3: Attention,* ed. A. Sanders & O. Neumann. Academic Press. [IB]
Emmorey, K., Corina, D. & Bellugi, U. (1995) Differential processing of topographic and referential functions of space. In: *Language, gesture and space,* ed. K. Emmorey & J. Reilly. Erlbaum. [BW]
Emmorey, K., Damasio, H., McCullough, S., Grabowski, T., Ponto, L. L. B., Hichwa, R. D. & Bellugi, U. (2002) Neural systems underlying spatial language in American Sign Language. *NeuroImage* 17:812–24. [BW]
Evans, G. (1982) *Varieties of reference.* Clarendon. [GG]
Fodor, J. A. (1975) *The language of thought.* Crowell. [MP-P]
(1983) *The modularity of mind: An essay on faculty psychology.* MIT Press. [aJRH]
(1987) *Psychosemantics.* Bradford Books/MIT Press. [MP-P]
Fodor, J. & Lepore, E. (2002) *The compositionality papers.* Oxford University Press. [MP-P]
Fogassi, L. (2003) Action recognition in the primate motor cortex can be a basis for communication. Paper presented at the conference, *Vocalize to localize: A missing piece in the puzzling route towards language.* Grenoble, France, January 2003. [PFM]
Frangiskakis, J. M., Ewart, A. K., Morris, C. A., Mervis, C. B., Bertrand, J., Robinson, B. F., Klein, B. P., Ensing, G. J., Everett, L. A., Green, E. D., Proschel, C., Gutowski, N. J., Noble, M., Atkinson, D. L., Odelberg, S .J. & Keating, M. T. (1996) Lim-kinase 1 hemizygosity implicated in impaired visuospatial constructive cognition. *Cell* 86:59–69. [aJRH]
Franz, V. H., Gegenfurtner, K. R., Bülthoff, H. H. & Fahle, M. (2000) Grasping visual illusions: No evidence for a dissociation between perception and action. *Psychological Science* 11(1):20–25. [aJRH]
Freyd, J. J. (1987) Dynamic mental representation. *Psychological Review* 94:427–38. [IB]
Galati, G., Lobel, E., Vallar, G., Berthoz, A., Pizzamiglio, L. & Bihan, D. L. (2000) The neural basis of egocentric and allocentric coding of space in humans: A functional magnetic resonance study. *Experimental Brain Research* 133:156–64. [aJRH]

Galletti, C. & Battaglini, P. P. (1989) Gaze-dependent visual neurons in area V3A of monkey prestriate cortex. *Journal of Neuroscience* 9:1112–25. [aJRH]
Gaymard, B., Siegler, I., Rivaud-Pechoux, S., Israel, I., Pierrot-Deseilligny, C. & Berthoz, A. (2000) A common mechanism for the control of eye and head movements in humans. *Annals of Neurology* 47(6):819–22. [aJRH]
Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (2002) *Cognitive neuroscience.* Norton. [PFM]
Gelman, R. & Gallistel, C. R. (1978) *The child's understanding of number.* Academic Press. [aJRH]
Gentilucci, M., Scandolara, C. Pigarev, I. N. & Rizzolatti, G. (1983) Visual responses in the postarcuate cortex (area 6) of the monkey that are independent of eye position. *Experimental Brain Research* 50:464–68. [aJRH]
Gerardin, E., Sirigu, A., Lehericy, S., Poline, J. P., Gaymard, B., Marsault, C., Agid, Y. & Le Bihan, D. (2000) Partially overlapping neural networks for real and imagined hand movements. *Cerebral Cortex* 10:1093–1104. [BW]
Ghadakpour, L. (2003) Le systeme conceptuel, a l'interface entre le langage, le raisonnement et l'espace qualitatif: Vers un modele de representations ephemeres. (Unpublished doctoral dissertation.) Ecole Polytechnique, Paris. [J-LD]
Gibson, B. S. (1994) Visual attention and objects: One versus two or convex versus concave? *Journal of Experimental Psychology: Human Perception and Performance* 20(1):203–207. [aJRH]
Gibson, E. J. & Spelke, E. S. (1983) The development of perception. In: *Handbook of child psychology: Cognitive development, vol. 3,* ed. P. Mussen, J. H. Flavell & E. M. Markman. Jossey Bass. [SL]
Gibson, J. (1979) *The ecological approach to visual perception.* Houghton Mifflin. [GG]
Gillett, G. (1992) *Representation, meaning and thought.* Clarendon. [GG]
Givon, T. (1995) *Functionalism and grammar.* Benjamins. [aJRH]
Goel, V., Gold, B., Kapur, S. & Houle, S. (1998) Neuroanatomical correlates of human reasoning. *Journal of Cognitive Neuroscience* 10:293–302. [BW]
Goldberg, A. (1995) *Constructions: A construction grammar approach to argument structure.* University of Chicago Press. [MP-P]
Goldberg, G. (1985) Supplementary motor area structure and function: Review and hypothesis. *Behavioral and Brain Sciences* 8:567–616. [PFM]
Goodale, M. A., Meenan, J. P., Bülthoff, H. H., Nicolle, D. A., Murphy, K. J. & Racicot, C. I. (1994) Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology* 4(7):604–10. [MLA, IB, aJRH]
Goodale, M. A. & Milner, D. (1992) Separate visual pathways for perception and action. *Trends in Neurosciences* 15(1):20–25. [aJRH]
Gottlieb, J. & Goldberg, M. (1998) The representation of visual salience in monkey parietal cortex. *Nature* 391(29):481–84. [IB]
Gould, S. J. & Eldredge, N. (1977) Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology* 3:115–51. [J-LD]
Graziano, M. S. A., Hu, X. T. A. & Gross, C. G. (1997) Visuospatial properties of ventral premotor cortex. *Journal of Neurophysiology* 77(5):2268–92. [aJRH]
Graziano, M. S. A., Reiss, L. A. J. & Gross, C. G. (1999) A neuronal representation of the location of nearby sounds. *Nature* 397(6718):428–30. [aJRH]
Gross, C. G. (1992) Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B - Biological Sciences* 335(1273):3–10. [aJRH]
Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G. & Blake, R. (2000) Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience* 12(5):711–20. [AK]
Haack, S. (1994) Deviant logics. In: *The encyclopedia of language and linguistics,* ed. R. E. Asher & J. M. Y. Simpson. Pergamon Press. [rJRH]
Hale, K. & Keyser, S. J. (2002) *Prolegomenon to a theory of argument structure.* MIT Press. [MP-P]
Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B. & Rapoport, S. I. (1991) Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences USA* 88:1621–25. [aJRH]
Hendry, S. H. C. & Yoshioka, T. (1994) A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus. *Science* 264:575–77. [aJRH]
Heywood, C. A. & Kentridge, R. W (2000) Affective blindsight? *Trends in Cognitive Sciences* 4(4):125–26. [aJRH]
Higginbotham, J. T. (1985) On semantics. *L.I.* 16(4):547–93. [MP-P]
(1989) Elucidations of meaning. *Linguistic and Philosophy* 12(3):465–517. [MP-P]
Hinton, G. E., Plaut, D. C. & Shallice, T. (1993) Simulating brain damage. *Scientific American,* 269(4):58–65. [GVJ]
Hockett, C. F. (1958) *A course in modern linguistics.* Macmillan. [GVJ]
Holland, J. H. (1992) *Adaptation in natural and artificial systems.* MIT Press. [J-LD]
Hubel, D. H. & Wiesel, T. N. (1968) Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology* 195:215–43. [aJRH]
Hume, D. (1740/1978) *A treatise of human nature.* Clarendon Press. [FC]
Hummel, J. (1999) Binding problem. In: *The MIT encyclopedia of the cognitive sciences,* ed. R. A. Wilson & F. C. Keil. MIT Press. [aJRH]
Hurford, J. R. (1987) *Language and number: The emergence of a cognitive system.* Basil Blackwell. [aJRH]
(1999) Individuals are abstractions. *Behavioral and Brain Sciences* 22(4):620–21. [aJRH]
(2000a) Languages treat 1–4 specially. *Mind and Language* 16(1):69–75. [aJRH]
(2000b) Social transmission favours linguistic generalization. In: *The evolutionary emergence of language: Social function and the origins of linguistic form,* ed. C. Knight, M. Studdert-Kennedy & J. Hurford. Cambridge University Press. [aJRH]
(2001) Protothought had no logical names. In: *New essays on the origin of language,* ed. J. Trabant & S. Ward. De Gruyter. [aJRH]
(2002) The roles of expression and representation in language evolution. In: *The transition to language,* ed. A. Wray. Oxford University Press. [aJRH, AK]
Hutchins, E. (1995) *Cognition in the wild.* MIT Press. [IB]
Indurkhya, B. (1992) *Metaphor and cognition.* Kluwer Academic. [BI]
(1998) On creation of features and change of representation. *Journal of Japanese Cognitive Science Society* 5(2):43–56. [BI]
Ingle, D. J. (1973) Two visual systems in the frog. *Science* 181:1053–55. [aJRH]
(1980) Some effects of pretectum lesions on the frog's detection of stationary objects. *Behavioral Brain Research* 1:139–63. [aJRH]
(1982) Organization of visuomotor behaviors in vertebrates. In: *Analysis of visual behavior,* ed. D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield. MIT Press. [aJRH]
Jackendoff, R. (2002) *Foundations of language.* Oxford University Press. [MP-P]
Jackendoff, R. & Landau, B. (1992) Spatial language and spatial cognition. In: *Languages of the mind: Essays on mental representation,* ed. R. Jackendoff. MIT Press. [aJRH]
Jeannerod, M., Arbib, M. A., Rizzolatti, G. & Sakata, H. (1995) Grasping objects – the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences* 18:(7)314–20. [aJRH]
Johansson, G. (1973) Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics* 14:201–11. [AK]
Johnson, A. (2003) *Families of the forest: The Matsigenka Indians of the Peruvian Amazon.* University of California Press. [aJRH]
Johnson-Laird, P. N. & Steedman, M. (1978) The psychology of syllogisms. *Cognitive Psychology* 10:64–99. [GP]
Johnsrude, I. S., Owen, A. M., Crane, J., Milner, B. & Evans, A. C. (1999) A cognitive activation study of memory for spatial relationships. *Neuropsychologia* 37:829–41. [aJRH]
Jones, G. V. (1985) Deep dyslexia, imageability, and ease of predication. *Brain and Language* 24:1–19. [GVJ]
(2002) Predicability (ease of predication) as semantic substrate of imageability in reading and retrieval. *Brain and Language* 82:159–66. [GVJ]
Kaas, J. H. & Hackett, T. A. (1999) "What" and "where" processing in auditory cortex. *Nature Neuroscience* 2(12):1045–47. [aJRH]
Kahneman, D. & Treisman. A. (1984) Changing views of attention and automaticity. In *Varieties of attention,* ed. R. Parasuraman & D. A. Davies. Academic Press. [aJRH]
(1992) The reviewing of object files: Object-specific integration of information. *Cognitive Psychology* 24:175–219. [aJRH]
Kamp, H. (1981) A theory of truth and semantic representation. In: *Formal methods in the study of language: Proceedings of the Third Amsterdam Colloquium, Mathematical Centre Tracts, pp. 277–322,* ed. J. Groenendijk, T. Janssen & M. Stokhof. [Reprinted in: *Truth, interpretation and information* (1984), ed. J. Groenendijk, T. M. V. Janssen & M. Stokhof.] Foris. [rJRH]
Kamp, H. & Partee, B. (1995) Prototype theory and compositionality. *Cognition,* 57(2):129–91. [J-LD]
Kamp, H. & Reyle, U. (1993) *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory.* Kluwer Academic. [arJRH]
Kant, I. (1783) *Prolegomena zu einer jeden künftigen Metaphysik die als Wissenschaft wird auftreten können.* Johann Friedrich Hartknoch. [aJRH]
(1789/1929) *Critique of pure reason,* trans. N. Kemp-Smith. Macmillan. [GG]
(1905) *Prolegomena to any future metaphysics,* trans. P. Carus. Open Court. [aJRH]
Keenan, E. (1987) Facing the truth. In: *Universal grammar: Fifteen essays,* ed. E. Keenan. Croom Helm. [aJRH]
Kikuchi-Yorioka, Y. & Sawaguchi, T. (2000) Parallel visuospatial and audiospatial working memory processes in the monkey dorsolateral prefrontal cortex. *Nature Neuroscience* 3(11):1075–76. [aJRH]
Kirby, S. (1999) Syntax out of learning: The cultural evolution of structured communication in a population of induction algorithms. In: *Advances in*

artificial life, ed. D. Floreano, J. D. Nicoud & F. Mondada. *Lecture Notes in Computer Science,* No. 1674. Springer. [aJRH]
(2000) Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In: *The evolutionary emergence of language: Social function and the origins of linguistic form,* ed. C. Knight, M. Studdert-Kennedy & J. R. Hurford. Cambridge University Press. [J-LD, aJRH]
Korte, M. & Rauschecker, J. P. (1993) Auditory spatial tuning of cortical neurons is sharpened in cats with early blindness. *Journal of Neurophysiology* 70(4):1717–21. [aJRH]
Kosslyn, S. M., DiGirolamo, G. J., Thompson, W. L. & Alpert, N. M. (1998) Mental rotation of objects versus hands: Neural mechanisms revealed by positron emission tomography. *Psychophysiology* 35:151–61. [BW]
Kotovsky, A. & Baillargeon, R. (1998) The development of calibration-based reasoning about collision events in young infants. *Cognition* 67:311–51. [PFD]
Krebs, J. R. & Dawkins, R. (1984) Animal signals: Mind-reading and manipulation. In: *Behavioural ecology: An evolutionary approach,* ed. J. R. Krebs & N. B. Davies. Blackwell Scientific. [aJRH]
Kripke, S. (1980) *Naming and necessity.* Blackwell. [aJRH]
Landau, B. & Jackendoff, R. (1993) "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16(2):217–38. [BB, arJRH]
Landau, B, Smith, L. B. & Jones, S. S. (1988) The importance of shape in early lexical learning. *Cognitive Development* 3:299–321. [aJRH]
(1998a) Object perception and object naming in early development. *Trends in Cognitive Sciences* 2(1):19–24. [aJRH]
(1998b) Object shape, object function and object name. *Journal of Memory and Language* 38:1–27. [aJRH]
Langer, S. K. (1957) *Philosophy in a new key, 3rd edition.* Harvard University Press. [BI]
Lehky, S. R. & Sejnowski, T. J. (1988) Network model of shape-from-shading – neural function arises from both receptive and projective fields. *Nature* 333:452–54. [aJRH]
Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. & Pitts, W. H. (1959) What the frog's eye tells the frog's brain. *Proceedings of the Institute of Radio Engineers* 47:1940–51. [Reprinted in McCulloch 1988]. [MP-P]
Liddell, S. K. (1990) Four functions of a locus: Re-examining the structure of a space in ASL. In: *Sign language research: Theoretical issues,* ed. C. Lucas. Gallaudet University Press. [aJRH]
Livingstone, M. & Hubel, D. (1988) Segregation of form, color, movement, and depth: anatomy, physiology and perception. *Science* 240:740–69. [aJRH]
Locke, J. (1689/1975) *An essay concerning human understanding.* Clarendon Press. [FC]
Luck, S. J. & Vogel, E. K. (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390:279–81. [aJRH]
Luria, A. R. (1973) *The working brain.* Penguin. [GG]
Luschei, E. S. & Goldberg, L. J. (1981) Neural mechanisms of mandibular control: Mastication and voluntary biting. In: *Handbook of physiology: The nervous system, vol. 2,* ed. J. M. Brookhart. American Physiological Society. [PFM]
Lyons, J. (1977) *Semantics (2 vols.).* Cambridge University Press. [aJRH]
MacNeilage, P. F. (1998) The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21:499–546. [PFM]
MacNeilage, P. F. & Davis, B. L. (2000) On the origin of internal structure of word forms. *Science* 288:527–31. [PFM]
(2001) Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology* 11:696–700. [PFM]
MacSweeney, M., Woll, B., Campbell, R., Calvert, G. A., McGuire, P. K., David, A. S., Simmons, A. & Brammer, M. J. (2002) Neural correlates of British Sign Language processing: Specific regions for topographic language? *Journal of Cognitive Neuroscience* 14(7):1064–75. [BW]
Mandler, G. & Shebo, B. J (1982) Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General* 111:1–22. [aJRH]
Marcel, A. J. (1998) Blindsight and shape perception: Deficit of visual consciousness or of visual function? *Brain* 121:1565–88. [aJRH]
Marois, R., Chung Leung, H. & Gore, J. (2000) A stimulus-driven approach to object identity and location processing in the human brain. *Neuron* 25:717–28. [aJRH]
Martin, A., Wiggs, C. L, Ungerleider, L. G. & Haxby, J. V. (1996) Neural correlates of category-specific knowledge. *Nature* 379:649–52. [aJRH]
Maye, A. (2002) Neuronale Synchronität, zeitliche Bindung und Wahrnehmung. (Unpublished doctoral dissertation). TU. [MW]
Mayr, E. (1982) *The growth of biological thought.* Belknap. [PFM]
McCarthy, J. (1968) Programs with common sense. In: *Semantic information processing,* ed. M. L. Minsky. MIT Press. [TJA]
McCulloch, W. S. (1988) *Embodiments of mind.* MIT Press. [MP-P]
McCulloch, W. S. & Pitts, W. H. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5:115–33. [TJA]
McDonald, B. H. (1994) Sign language morphology. In: *The encyclopedia of language and linguistics,* ed. R. E. Asher & J. M. Y. Simpson. Pergamon Press. [aJRH]
Meredith, M. A. & Stein, B. E. (1983) Interactions among converging sensory inputs in the superior colliculus. *Science* 221:389–91. [TJA]
(1986) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology* 56:640–22. [TJA]
Merigan, W. H, Byrne, C. E. & Maunsell, J. H. R. (1991) Does primate motion perception depend on the magnocellular pathway? *The Journal of Neuroscience* 11(11):3422–29. [aJRH]
Merigan, W. H. & Maunsell, J. H. R. (1993) How parallel are the primate visual pathways? *Annual Review of Neuroscience* 16:369–402. [aJRH, MW]
Michalski, R. S. & Stepp, R. E. (1983) Learning from observation: Conceptual clustering. In: *Machine learning – An artificial approach.* Tioga. [J-LD]
Michotte, A. E. (1946/1963) *The perception of causality,* trans. T. R. Miles & E. Mile. Basic Books. [SL]
Milner, A. D. (1998) Streams and consciousness: Visual awareness and the brain. *Trends in Cognitive Sciences* 2(1):25–30. [aJRH]
Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action.* Oxford University Press. [BB, aJRH, PFM]
Mishkin, M., Ungerleider, L. G. & Macko, K. A. (1983) Object vision and spatial vision: two cortical pathways. *Trends in Neuroscience* 6:414–17. [aJRH]
Mithun, M. (2000) Noun and verb in Iroquoian languages: Multicategorisation from multiple criteria. In: *Approaches to the typology of word classes,* ed. P. M. Vogel & B. Comrie. Mouton de Gruyter. [MW]
Montague, R. (1970) English as a formal language. In: *Linguaggi nella Società e nella Tecnica,* ed. B.Visentini. Edizioni di Comunità. [aJRH]
(1973) The proper treatment of quantification in ordinary English. In: *Approaches to natural language,* ed. J. Hintikka, J. Moravcsik & P. Suppes. Reidel. [aJRH]
Moore, T., Rodman, H. R. & Gross, C. G. (1998) Man, monkey, and blindsight. *Neuroscientist* 4(4):227–30. [aJRH]
Morris, C. W. (1938) *Foundations of the theory of signs.* Chicago University Press. [aJRH]
(1946) *Signs, language, and behavior.* Prentice-Hall. [aJRH]
Morris, J. S, Ohman, A. & Dolan, R. J. (1999) A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences of the USA* 96(4):1680–85. [aJRH]
Murdock, G. P. (1960) *Social structure.* Macmillan. [aJRH]
Nakamura, K. (1999) Auditory spatial discriminatory and mnemonic neurons in rat posterior parietal cortex. *Journal of Neurophysiology* 82(5):2503–17. [aJRH]
Neisser, U. (1967) *Cognitive psychology.* Appleton-Century-Crofts. [aJRH]
Newell, A. (1982) The knowledge level. *Artificial Intelligence* 18:82–127. [TJA]
Newman, E. A. & Hartline, P. H. (1981) Integration of visual and infrared information in bimodal neurons of the rattlesnake optic tectum. *Science* 213:789–91. [TJA]
Nowak, M. A., Plotkin, J. B & Jansen, V. A. A. (2000) The evolution of syntactic communication. *Nature* 404:495–98. [aJRH]
O'Brien, G. & Opie, J. (1999) A connectionist theory of phenomenal experience. *Behavioral and Brain Sciences* 22(1):127–48. [aJRH]
O'Craven, K., Downing, P. & Kanwisher, N. (1999) fMRI evidence for objects as the units of attentional selection. *Nature* 401(7):584–87. [IB]
Padden, C. (1990) The relation between space and grammar in ASL verb morphology. In: *Sign language research: Theoretical issues,* ed. C. Lucas. Gallaudet University Press. [aJRH]
Paivio, A. (1971) *Imagery and verbal processes.* Holt, Rinehart and Winston. [GVJ]
(1983) The empirical case for dual coding. In: *Imagery, memory and cognition,* ed. J. C. Yuille. Erlbaum. [GVJ]
Parsons, T. (1990) *Events in the semantics of English: A study in subatomic semantics.* MIT Press. [aJRH]
Patton, P. E. & Anastasio, T. J. (2003) Modeling cross-modal enhancement and modality-specific suppression in multisensory neurons. *Neural Computation* 15:783–810. [TJA]
Pepperberg, I. (1999) Can a parrot respond to absence of information? In: *The Alex studies,* ed. I. Pepperberg. Harvard University Press. [rJRH]
Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K. & Ortega, J. E. (1989) Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology* 146:87–113. [aJRH]
Piaget, J. (1953) *The origin of intelligence in the child,* trans. M. Cook. Routledge. (Also published Penguin, 1977.) [BI]
(1962) *Play, dreams and imitation in childhood,* trans. C. Gattegno & F. M. Hodgson. Norton. [BI]
Pianesi, F. & Varzi, A. C. (2000) Events and event talk. In: *Speaking of events,* ed. J. Higginbotham, F. Pianesi & A. C. Varzi. Oxford University Press. [MW]
Plaut, D. C. (1999) A connectionist approach to word reading and acquired

dyslexia: Extension to sequential processing. *Cognitive Science* 23:543–68. [GVJ]
Plaut, D. C. & Shallice, T. (1993) Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology* 10:377–500. [GVJ]
Pustejovsky, J. (1995) *The generative lexicon.* MIT Press. [MP-P]
Pylyshyn, Z. W. (1989) The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition* 32:65–97. [aJRH]
(1999) Is vision continuous with cognition? – The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences* 22(3):341–423. [IB]
(2000) Situating vision in the world. *Trends in Cognitive Sciences* 4(5):197–207. [aJRH, IB]
Quine, W. V. (1961) On what there is. In: Quine, M. V. (1961) *From a logical point of view, 2nd edition.* Harvard University Press. [MW]
Ramachandran, V. S. & Blakeslee, S. (1998) *Phantoms in the brain: Human nature and the architecture of the mind.* Fourth Estate. [aJRH]
Rauschecker, J. P. (1997) Processing of complex sounds in the auditory cortex of cat, monkey, and man. *Acta Otolaryngolica Supplement* 532:34–38. [aJRH]
Redican, W. K. (1975) Facial expressions in nonhuman primates. In: *Primate behavior: Developments in field and laboratory research, vol. 4,* ed. L. A. Rosenblum. Academic Press. [PFM]
Reichle, E. D., Carpenter, P. & Just, M. A. (2000) The neural bases of strategy and skill in sentence-picture verification. *Cognitive Psychology* 40:261–95. [BW]
Rizzolatti, G. & Arbib, M. A. (1998) Language within our grasp. *Trends in Neuroscience* 21(5):188–94. [MAA, BB, PFD, aJRH, PFM]
Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S. & Rauschecker, J. P. (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience* 2(12):1131–36. [aJRH]
Russac, R. J. (1983) Early discrimination among small object collections. *Journal of Experimental Child Psychology* 36:124–38. [aJRH]
Russell, B. (1905) On denoting. *Mind* 14:479–93. [aJRH]
(1912/1998) *The problems of philosophy.* Oxford University Press. [FC]
(1918/1985) *The philosophy of logical atomism.* Open Court. [FC]
(1940) *An inquiry into meaning and truth.* Allen and Unwin. [aJRH]
(1948) *Human knowledge, its scope and limits.* Allen and Unwin. [aJRH]
(1957) Mr. Strawson on referring. *Mind* 66:385–89. [aJRH]
(1959) *My philosophical development.* Allen and Unwin. [aJRH]
Ryle, G. (1957) The theory of meaning. In: *British philosophy in the mid-century,* ed. C. A. Mace. Allen and Unwin. [aJRH]
Sanders, M. D., Warrington, E. K., Marshall, J. & Weiskrantz, L. (1974) "Blindsight": Vision in a field defect. *Lancet* 20:707–708. [aJRH]
Schaeffer, B., Eggleston, V. H. & Scott, J. L. (1974) Number development in young children. *Cognitive Psychology* 6:357–79. [aJRH]
Schein, B. (1993) *Plurals and events.* MIT Press. [MP-P]
Schillen, T. B. & König, P. (1994) Binding by temporal structure in multiple feature domains of an oscillatory neuronal network. *Biological Cybernetics* 70:397–405. [MW]
Scholl, B. J. & Nakayama, K. (2002) Causal capture: Contextual effects on the perception of collision events. *Psychological Science* 13:493–98. [SL]
Searle, J. R. (1979) *Expression and meaning: Studies in the theory of speech acts.* Cambridge University Press. [aJRH]
Singer, W. & Gray, C. M. (1995) Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience* 18:555–86. [MW]
Smith, L. B., Jones, S. S. & Landau, B. (1996) Naming in young children: A dumb attentional mechanism? *Cognition* 60:143–171. [aJRH]
Snell, W. (1964) *Kinship relations in Machiguenga.* Master's thesis, Hartford Seminary. Hartford, CT. [aJRH]
Stanner, W. E. H. (1937) Aboriginal modes of address and reference in the northwest of the Northern Territory. *Oceania* 7:300–15. [aJRH]
Starkey, P. & Cooper, R. G. (1980) Perception of numbers by human infants. *Science* 210:1033–35. [aJRH]
Steels, L. (1997) Constructing and sharing perceptual distinctions. Coupling of meaning creation and naming games in software experiments. In: *Proceedings of the European conference on machine learning,* ed. M. van Someren & G. Widmer. Springer. [MLA]
Stein, B. E. & Meredith, M. A. (1993) *The merging of the senses.* MIT Press. [TJA]
Stein, B. E. & Wallace, M. T. (1996) Comparisons of cross-modality integration in midbrain and cortex. *Progress in Brain Research* 112:289–99. [TJA]
Strawson, P. F. (1950) On referring. *Mind* 59:320–44. [aJRH]
(1959) *Subject and predicate in logic and grammar.* Methuen. [aJRH]
(1974) *Individuals: An essay in descriptive metaphysics.* Methuen. [aJRH]
Stricanne, B., Andersen, R. A & Mazzoni, P. (1996) Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *Journal of Neurophysiology* 76(3):2071–76. [aJRH]
Talmy, L. (1975) Semantics and syntax of motion. In: *Syntax and semantics, vol.4,* ed. J. P. Kimball. Academic Press. [SL]
(2000) *Toward a cognitive semantics.* MIT Press. [rJRH, SL]
Thornton, I., Rensink, R. & Shiffrar, M. (2002) Active vs. passive processing of biological motion. *Perception* 31:837–53. [AK]
Tian, B. & Rauschecker, J. P. (1998) Processing of frequency-modulated sounds in the cat's posterior auditory field. *Journal of Neurophysiology* 79(5):2629–42. [aJRH]
Trevarthen, C. B. (1968) Two mechanisms of vision in primates. *Psychologische Forschung* 31:299–337. [aJRH]
(2001) The neurobiology of early communication: Intersubjective regulations in human brain development. *Handbook of Brain and Behaviour in Human Development* 841–81. [GG]
Trick, L. M. & Pylyshyn, Z. W. (1993) What enumeration studies tell us about spatial attention: Evidence for limited capacity pre-attentive processing. *Journal of Experimental Psychology – Human Perception and Performance* 19(2):331–51. [aJRH]
(1994) Why are small and large numbers enumerated differently? A limited capacity pre-attentive stage in vision. *Psychological Review* 10:1–23. [aJRH]
Ullman, S. (1984) Visual routines. *Cognition* 18:97–157. [aJRH]
Ungerleider, L. G., Courtney, S. M. & Haxby, J. V. (1998) A neural system for human visual working memory. *Proceedings of the National Academy of Sciences USA* 95:883–90. [aJRH]
Ungerleider, L. G. & Mishkin, M. (1982) Two cortical visual systems. In: *Analysis of visual behavior,* ed. D. J. Ingle, M. A.Goodale, & R. J. W., Mansfield. MIT Press. [aJRH]
Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P. & Belliveau, J. W. (2001) Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences USA* 98(20):11656–61. [rJRH]
Valdes-Sosa, M., Cobo, A. & Pinilla, T. (1998) Transparent motion and object-based attention. *Cognition* 66:B13–B23. [aJRH]
Valiant, L. G. (1994) *Circuits of the mind.* Oxford University Press. [TJA]
Van Essen, D. C, Anderson, C. H. & Felleman, D. J. (1992) Information processing in the primate visual system: An integrated systems perspective. *Science* 255:419–23. [aJRH]
Vecera, S. P. & Farah, M. J. (1994) Does visual attention select objects or locations? *Journal of Experimental Psychology: General* 123(2):146–60. [aJRH]
Walker, S. (1983) *Animal thought.* Routledge. [aJRH]
Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, C. M., Cohen, L. G., Hallett, M. & Rauschecker, J. P. (2000) A positron emission tomographic study of auditory localization in the congenitally blind. *Journal of Neuroscience* 20(7):2664–72. [aJRH]
Wegener, P. (1885/1971) The life of speech. In: *Speech and reason,* ed. P. Wegener & D. W. Abse. University of Virginia Press. [BI]
Weiskrantz, L. (1986) *Blindsight: A case study and implications.* Oxford University Press. [aJRH]
(1997) *Consciousness lost and found.* Oxford University Press. [aJRH, IB]
Weiskrantz, L., Warrington, E., Sanders, M. & Marshall, J. (1974) Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain* 97:709–28. [aJRH]
Werning, M. (2001) How to solve the problem of compositionality by oscillatory networks. In: *Proceedings of the twenty-third annual conference of the cognitive science society,* ed. J. D. Moore & K. Stenning. Erlbaum. [MW]
(2003a) Synchrony and composition: Toward a cognitive architecture between classicism and connectionism. In: *Applications of mathematical logic in philosophy and linguistics,* ed. B. Loewe, W. Malzkorn & T. Raesch. Kluwer. [MW]
(2003b) *The temporal dimension of thought: Cortical foundations of predicative representation.* Synthese. (in press). [MW]
Westwood, D. A., Chapman, C. D. & Roy, E. A. (2000) Pantomimed actions may be controlled by the ventral visual stream. *Experimental Brain Research* 130:545–48. [aJRH]
Williams, J. M. G., Healy, H. G. & Ellis, N. C. (1999) The effect of imageability and predicability of cues in autobiographical memory. *Quarterly Journal of Experimental Psychology* 52A:555–79. [GVJ]
Wittgenstein, L. (1922) *Tractatus logico-philosophicus.* Routledge. [FC]
Woll, B. & Kyle, J. (1994) Sign language. In: *The encyclopedia of language and linguistics,* ed. R. E. Asher & J. M. Y. Simpson. Pergamon Press. [aJRH]
Wray, A. (1998) Protolanguage as a holistic system for social interaction. *Language and Communication* 18:47–67. [BI]
(2000) A protolanguage with no declaratives and no names. *Proceedings of the 3rd International Conference on the Evolution of Language, Paris.* [BI]
Wray, A. & Perkins, M. R. (2000) The functions of formulaic language: An integrated model. *Language and Communication* 20:1–28. [BI]
Wurtz, R. H. & Kandel, E. R. (2000) Perception of motion, depth, and form. In: *Principles of neural science,* ed. E. R. Kandel, J. H. Schwartz & T. M. Jessell. McGraw-Hill. [MW]
Yantis, S. (1998) Control of visual attention. In: *Attention,* ed. H. Pashler. Psychology Press. [IB]

Yantis, S. & Johnson, D. N. (1990) Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance* 16:812–25. [IB]

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L. & Raichle, M. E. (2001a) Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience* 4:651–55. [SL]

Zacks, J. M. & Tversky, B. (2001) Event structure in perception and conception. *Psychological Bulletin* 127:3–21. [SL]

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

Zadeh, L. A. & Kacprzyk, J., eds. (1992) *Fuzzy logic for the management of uncertainty.* Wiley. [rJRH]

Zeki, S. (1993) *A vision of the brain.* Blackwell Scientific. [aJRH]

Zhuo, Y., Zhou, T. G., Rao, H. Y., Wang, J. J., Meng, M., Chen, M., Zhou, C. & Chen, L. (2003) Contributions of the visual ventral pathway to long-range apparent motion. *Science* 299(5605):417–20. [rJRH]