BEHAVIORAL AND BRAIN SCIENCES (2002) **25,** 657–726
*Printed in the United States of America*

# The cognitive functions of language

**Peter Carruthers**
*Department of Philosophy, University of Maryland, College Park, MD 20742*
**pcarruth@umd.edu**
**www.philosophy.umd.edu/people/faculty/pcarruthers**

**Abstract:** This paper explores a variety of different versions of the thesis that natural language is involved in human thinking. It distinguishes amongst strong and weak forms of this thesis, dismissing some as implausibly strong and others as uninterestingly weak. Strong forms dismissed include the view that language is conceptually necessary for thought (endorsed by many philosophers) and the view that language is de facto the medium of all human conceptual thinking (endorsed by many philosophers and social scientists). Weak forms include the view that language is necessary for the acquisition of many human concepts and the view that language can serve to scaffold human thought processes. The paper also discusses the thesis that language may be the medium of *conscious* propositional thinking, but argues that this cannot be its most fundamental cognitive role. The idea is then proposed that natural language is the medium for non-domain-specific thinking, serving to integrate the outputs of a variety of domain-specific conceptual faculties (or central-cognitive "quasi-modules"). Recent experimental evidence in support of this idea is reviewed and the implications of the idea are discussed, especially for our conception of the architecture of human cognition. Finally, some further kinds of evidence which might serve to corroborate or refute the hypothesis are mentioned. The overall goal of the paper is to review a wide variety of accounts of the cognitive function of natural language, integrating a number of different kinds of evidence and theoretical consideration in order to propose and elaborate the most plausible candidate.

**Keywords:** Cognitive evolution; conceptual module; consciousness; domain-general; inner speech; language; logical form (LF); thought

## 1. Introduction

Natural language looms large in the cognitive lives of ordinary folk. Although proportions vary, many people seem to spend a good deal of their waking activity engaged in "inner speech," with imaged natural language sentences occupying a significant proportion of the stream of their conscious mentality.

This bit of folk wisdom has been corroborated by Hurlburt (1990; 1993), who devised a method for sampling people's inner experience. Subjects wore headphones during the course of the day, through which they heard, at various intervals, a randomly generated series of bleeps. When they heard a bleep, they were instructed to immediately "freeze" what was passing through their consciousness at that exact moment and then make a note of it, before elaborating on it later in a follow-up interview. Although frequency varied widely, all normal (as opposed to schizophrenic) subjects reported experiencing inner speech on some occasions – with the minimum being 7 percent of occasions sampled, and the maximum being 80 percent. Most subjects reported inner speech on more than half of the occasions sampled. (The majority of subjects also reported the occurrence of visual images and emotional feelings – on between 0 percent and 50 percent of occasions sampled in each case). Think about it: more than half of the total set of moments which go to make up someone's conscious waking life occupied with inner speech – that's well-nigh continuous!

Admittedly, the sample-sizes in Hurlburt's studies were small; and other interpretations of the data are possible. (Perhaps the reports of linguistically-clothed thoughts occurring at the time of the bleep were a product of *confabulation,* for example, reflecting people's naïve theory that thought *must* be in natural language. If so, this should be testable.) But let us suppose that inner verbalization is as ubiquitous as common-sense belief and Hurlburt's data would suggest. Just what would all this inner verbalization be doing? What would be its function, or cognitive role? The naïve common-sense answer is that inner verbalization is constitutive of our thinking – it is that we think *by* talking to ourselves in inner speech (as well as by manipulating visual images, etc.). Anyone who holds such a view endorses a version of what I shall call "the cognitive conception of language," which maintains that, besides its obvious communicative functions, language also has a direct role to play in normal human cognition (in thinking and reasoning).

Quite a different answer would be returned by most members of the cognitive science community, however. They endorse what I shall call "the (purely) communicative conception of language," according to which language is but an input-output system for central cognition. Believing

Peter Carruthers is Professor of Philosophy and Chair of the Philosophy Department at the University of Maryland, College Park. Until recently he Directed the Hang Seng Centre for Cognitive Studies at the University of Sheffield. He is the author of eight books, coeditor of four, and has published numerous articles in journals and collections. There are two major foci to his main research area in philosophy of psychology. One is the nature of consciousness, where he has developed and defended a form of higher-order thought theory. The other is the architecture of human cognition, and the place of natural language within it.

 0140-525X/02 $12.50 

that language is only a channel, or *conduit,* for transferring thoughts into and out of the mind, they are then obliged to claim that the stream of inner verbalization is more-or-less epiphenomenal in character. (Some possible minor cognitive roles for inner speech, which should nevertheless be acceptable to those adopting this perspective, will be canvassed later.) The real thinking will be going on elsewhere, in some other medium of representation.

One reason for the popularity of the communicative conception amongst cognitive scientists is that almost all now believe that language is a distinct input-output module of the mind (at least in *some* sense of "module," if not quite in Fodor's classic – 1983 – sense). And they find it difficult to see how the language faculty could *both* have this status *and* be importantly implicated in central cognition. But this reasoning is fallacious. Compare the case of visual imagination: Almost everyone now thinks that the visual system is a distinct input-module of the mind, containing a good deal of innate structure. But equally, most cognitive scientists now accept that visual *imagination* re-deploys the resources of the visual module for purposes of reasoning – for example, many of the same areas of the visual cortex are active when imagining as when seeing. (For a review of the evidence, see Kosslyn 1994.)

What is apparent is that central cognition can co-opt the resources of peripheral modules, activating some of their representations to subserve central cognitive functions of thinking and reasoning. The same is then possible in connection with language. It is quite consistent with language being an innately structured input and output module, that central cognition should access and deploy the resources of that module when engaging in certain kinds of reasoning and problem solving.

Note, too, that hardly anyone is likely to maintain that visual imagery is a mere epiphenomenon of central cognitive reasoning processes, playing no real role in those processes in its own right. On the contrary, it seems likely that there are many tasks which we cannot easily solve without deploying a visual (or other) image. For example, suppose you are asked (orally) to describe the shape which is enclosed within the capital letter "A." It seems entirely plausible that success in this task should require the generation of a visual image of that letter, from which the answer ("a triangle") can then be read off. So it appears that central cognition operates, in part, by co-opting the resources of the visual system to generate visual representations, which can be of use in solving a variety of spatial-reasoning tasks. And this then opens up the very real possibility that central cognition may also deploy the resources of the language system to generate representations of natural language sentences (in "inner speech"), which can similarly be of use in a variety of *conceptual* reasoning tasks.

There is at least one further reason why the cognitive conception of language has had a bad press within the cognitive science community in recent decades. (It continues to be popular in some areas of the social sciences and humanities, including philosophy.) This is that many of the forms of the thesis which have been defended by philosophers and by social scientists are implausibly strong, as we shall see in section 3 below. The unacceptability of these strong views has then resulted in *all* forms of the cognitive conception being tarred with the same brush.

A crucial liberalizing move, therefore, is to realize that the cognitive conception of language can come in many different strengths, each one of which needs to be considered separately on its own merits. In this paper I shall distinguish between some of the many different versions of the cognitive conception. I shall begin (in sect. 2) by discussing some weak claims concerning the cognitive functions of language that are largely uncontroversial. This will help to clarify just what (any interesting form of) the cognitive conception *is,* by way of contrast. I shall then (in sect. 3) consider some claims which are so strong that cognitive scientists are clearly right in rejecting them, before zeroing in on those which are both interesting and plausible (in sects. 4 and 5). I shall come to focus, in particular, on the thesis that natural language is the medium of inter-modular integration. This is a theoretical idea which has now begun to gather independent empirical support. Finally (in sects. 6 and 7) some additional implications, elaborations, and possible further empirical tests of this idea are discussed.

I should explain at the outset, however, that the thesis I shall be working towards is that it is natural language *syntax* which is crucially necessary for inter-modular integration. The hypothesis is that non-domain-specific thinking operates by accessing and manipulating the representations of the language faculty. More specifically, the claim is that non-domain-specific thoughts implicate representations in what Chomsky (1995) calls "logical form" (LF). Where these representations are *only* in LF, the thoughts in question will be non-conscious ones. But where the LF representation is used to generate a full-blown phonological representation (an imaged sentence), the thought will generally be conscious.[1]

I should emphasize that I shall not be claiming that syntax is logically required for inter-modular integration, of course. Nor shall I be claiming that only natural language syntax – with its associated recursive and hierarchical structures, compositionality, and generativity – could possibly play such a role in any form of cognition, human or not. (In fact it is the phrase-structure element of syntax which does the work in my account; see sect. 6.1 below.) Rather, my claim will be that syntax *does* play this role in human beings. It is a factual claim about the way in which our cognition happens to be structured, not an unrestricted modal claim arrived at by some sort of task-analysis.

I should also declare at the outset how I shall be using the word "thought" in this paper. Unless I signal otherwise, I intend all references to *thought* and *thinking* to be construed realistically. Thoughts are discrete, semantically-evaluable, causally-effective states, possessing component structure, and where those structures bear systematic relations to the structures of other, related, thoughts. So distinct thoughts have distinct physical realizations, which may be true or false, and which cause other such thoughts and behavior. And thoughts are built up out of component parts, where those parts belong to types which can be shared with other thoughts. It is not presupposed, however, that thoughts are borne by sentence-like structures. Although I shall be arguing that *some* thoughts are carried by sentences (namely, non-domain-specific thoughts which are carried by sentences of natural language), others might be carried by mental models or mental images of various kinds.

It is hugely controversial that there *are* such things as thoughts, thus construed, of course. And while I shall say a little in defense of this assumption below (in sect. 3.3), for the most part it is just that – an assumption – for present purposes. I can only plead that one can't do everything in one article, and that one has to start somewhere. Those who don't want to share this assumption should read what follows conditionally: if we *were* to accept that there are such things as realistically-construed thoughts, then how, if at all, should they be seen as related to natural language sentences?

Also, a word about the nature of the exercise before we proceed further. This article ranges over a great many specialist topics and literatures in a number of distinct disciplines. Of necessity, therefore, our discussion of any given subject must be relatively superficial, with most of the detail, together with many of the required qualifications and *caveats,* being omitted. Similarly, my arguments against some of the competitor theories have to be extremely brisk, and some quite large assumptions will have to get taken on board without proper examination. My goal, here, is just to map out an hypothesis space, using quite broad strokes, and then to motivate and discuss what I take to be the most plausible proposal within it.

## 2. Weak claims

Everyone will allow that language makes some cognitive difference. For example, everyone accepts that a human being with language and a human being without language would be very different, cognitively speaking. In this section I shall outline some of the reasons why.

### 2.1. Language as the conduit of belief

Everyone should agree that natural language is a necessary condition for human beings to be capable of entertaining at least some kinds of thought. Language is the conduit through which we acquire many of our beliefs and concepts, and in many of these cases we could hardly have acquired the component concepts in any other way. Concepts which have emerged out of many years of collective labor by scientists, for example – such as *electron, neutrino,* and *DNA* – would de facto be inaccessible to someone deprived of language. This much, at any rate, should be obvious. But all it really shows is that language is required for certain kinds of thought; not that language is actually involved in or is the representational vehicle of those thoughts.

It is often remarked, too, that the linguistic and cognitive abilities of young children will normally develop together. If children's language is advanced, then so will be their abilities across a range of tasks; and if children's language is delayed, then so will be their cognitive capacities. To cite just one item from a wealth of empirical evidence: Astington (1996) and Peterson and Siegal (1998) report finding a high correlation between language-ability and children's capacity to pass false-belief tasks, whose solution requires them to attribute, and reason from, the false belief of another person. Does this and similar data show that language is actually involved in children's thinking?

In the same spirit, we might be tempted to cite the immense cognitive deficits that can be observed in those rare cases where children grow up without exposure to natural language. Consider, for example, the cases of so-called "wolf children," who have survived in the wild in the company of animals, or of children kept by their parents locked away from all human contact (Malson 1972; Curtiss 1977). Consider, also, the cognitive limitations of profoundly deaf children born of hearing parents, who have not yet learned to sign (Schaller, 1991; Sachs 1989). These examples might be thought to show that human cognition is constructed in such a way as to require the presence of natural language if it is to function properly.

But all that such data really show is, again, that language is a necessary condition for certain kinds of thought and types of cognitive process; not that it is actually implicated in those forms of thinking. And this is easily explicable from the standpoint of someone who endorses the standard cognitive science conception of language, as being but an input-output system for central cognition, or a mere communicative device. For language, in human beings, is a necessary condition of normal enculturation. Without language, there are many things which children cannot learn; and with delayed language, there are many things that children will learn only later. It is only to be expected, then, that cognitive and linguistic development should proceed in parallel. It does not follow that language is itself actually used in children's central cognition.

Another way of putting the point is that this proposed cognitive function of language is purely developmental – or diachronic – rather than synchronic. Nothing is said about the role of language in the cognition of adults, once a normal set of beliefs and concepts has been acquired. And the evidence from aphasia suggests that at least many aspects of cognition can continue to operate normally once language has been removed.

Aphasias come in many forms, of course, and in many different degrees of severity. And it is generally hard to know the extent of any collateral damage – that is, to know which other cognitive systems besides the language faculty may have been disabled as a result of the aphasia-causing brain-damage. But many patients with severe aphasia continue to be adept at visuo-spatial thinking, at least (Kertesz 1988), and many continue to manage quite well for themselves in their daily lives.

Consider, for example, the agrammatic aphasic man studied in detail by Varley (1998; 2002). He is incapable of either producing or comprehending sentences, and he also has considerable difficulty with vocabulary, particularly verbs. He has lost all mentalistic vocabulary ("belief," "wants," etc.), and his language system is essentially limited to nouns. Note that there is not a lot of explicit thinking that you can do using just nouns! (It should also be stressed that he has matching deficits of input and output, suggesting that it is the underlying system of linguistic knowledge which has been damaged.) Yet he continues to drive, and to have responsibility for the family finances. He is adept at communicating, using a mixture of single-word utterances and pantomime. And he has passed a range of tests of theory of mind (the standard battery of false-belief and deception tasks, explained using nouns and pantomime), as well as various tests of causal thinking and reasoning. It appears that, once language has done its developmental work of loading the mind with information, a good deal of adult cognition can thereafter survive its loss.

Since natural language is the conduit for many of our beliefs and for much of our enculturation, everyone should accept that language is immensely important for normal cognitive development. That language has this sort of cognitive function should be no news to anyone.[2]

### 2.2. Language as sculpting cognition

A stronger and more controversial thesis has been proposed and defended by some researchers over recent decades. This is that the process of language acquisition and enculturation does not merely serve to load the mind with beliefs and concepts, but actually sculpts our cognitive *processes* to some degree (Bowerman & Levinson 2001; Lucy 1992a; 1992b; Nelson 1996).[3]

For example, acquisition of Yucatec (as opposed to English) – in which plurals are rarely marked and many more nouns are treated grammatically as substance-terms like "mud" and "water" – leads subjects to see similarities amongst objects on the basis of material composition rather than shape (Lucy 1992b; Lucy & Gaskins 2001). And children brought up speaking Korean (as opposed to English) – in which verbs are highly inflected and massive noun ellipsis is permissible in informal speech – leads children to be much weaker at categorization tasks, but much better at means-ends tasks such as using a rake to pull a distant object towards them (Choi & Gopnik 1995; Gopnik 2001; Gopnik et al. 1996).

Fascinating as these data are, they do not, in themselves, support any version of the cognitive conception of language. This is because the reported effects of language on cognition are still entirely diachronic and developmental, rather than synchronic. The fact that acquiring one language as opposed to another causes subjects to attend to different things and to reason somewhat differently doesn't show that language itself is actually involved in people's thinking. Indeed, on the hypothesis proposed by Gopnik (2001), language-acquisition has these effects by providing evidence for a pre-linguistic theorizing capacity, which operates throughout development to construct children's systems of belief and inference.

### 2.3. Language as a cognitive scaffold

Other claims can be extracted from the work of Vygotsky (1934/1986), who argues that language and speech serve to scaffold the development of cognitive capacities in the growing child. Researchers working in this tradition have studied the self-directed verbalizations of young children – for example, observing the effects of their soliloquies on their behavior (Diaz & Berk 1992). They have found that children tend to verbalize more when task demands are greater, and that those who verbalize most tend to be more successful in problem-solving.

This claim of linguistic scaffolding of cognition admits, however, of a spectrum of readings. At its weakest, it says no more than has already been conceded above, that language may be a necessary condition for the acquisition of certain cognitive skills. At its strongest, on the other hand, the idea could be that language forms part of the functioning of the highest-level executive system – which would then make it a variant of the ideas to be discussed in sections 4 and 5 below.

Clark (1998) argues for a sort of intermediate-strength version of the Vygotskian idea, defending a conception of language as a cognitive *tool.* (Chomsky, too, has argued for an account of this sort. See his 1976, Ch. 2.) According to this view – which Clark labels "the supra-communicative conception of language" – certain extended processes of thinking and reasoning constitutively involve natural language. The idea is that language gets used, not just for communication, but also to augment human cognitive powers.

Thus by writing an idea down, for example, I can off-load the demands on memory, presenting myself with an object of further leisured reflection; and by performing arithmetic calculations on a piece of paper, I may be able to handle computational tasks which would otherwise be too much for me (and my short-term memory). In similar fashion, it may be that *inner* speech serves to enhance memory, since it is now well-established that the powers of human memory systems can be greatly extended by association (Baddeley 1988). Inner speech may thus facilitate complex trains of reasoning (Varley 1998).

Notice that according to this supra-communicative account, the involvement of language in thought only arises when we focus on a process of thinking or reasoning *extended over time.* So far as any given individual (token) thought goes, the account can (and does) buy into the standard input-output conception of language. It maintains that there is a neural episode which carries the content of the thought in question, where an episode of that type can exist in the absence of any natural language sentence and can have a causal role distinctive of the thought, but which in the case in question causes the production of a natural language representation. This representation can then have further benefits for the system of the sort which Clark explores (off-loading or enhancing memory).

According to stronger forms of the cognitive conception to be explored in later sections, in contrast, a particular tokening of an inner sentence is (sometimes) an inseparable part of the mental episode which carries the content of the thought-token in question. So there is no neural or mental event at the time which can exist distinct from that sentence, which can occupy a causal role distinctive of that sort of thought, and which carries the content in question; and so language is actually involved in (certain types of) cognition, even when our focus is on individual (token) thinkings.

In this section I have discussed two weak claims about the role of language (that language is necessary for the acquisition of many beliefs and concepts; and that language may serve as a cognitive tool, enhancing the range and complexity of our reasoning processes). These claims should be readily acceptable to most cognitive scientists. In addition, I have briefly introduced a more controversial thesis, namely that the acquisition of one or another natural language can sculpt our cognitive processes, to some degree. But this thesis relates only to the developmental, or diachronic, role of language. It says nothing about the role of language in adult cognition. We will in future focus on more challenging versions of the cognitive conception of language.

## 3. Strong claims

As is starting to emerge, the thesis that language has a cognitive function admits of a spectrum of readings. In this section I shall jump to the other end of that spectrum, considering forms of the cognitive conception of language which are too strong to be acceptable.

### 3.1. Language as necessarily required for thought

When the question of the place of natural language in cognition has been debated by philosophers the discussion has, almost always, been conducted a priori in universalist terms. Various arguments have been proposed for the claim that it is a conceptually necessary truth that *all* thought requires language, for example (Davidson 1975; 1982; Dummett 1981; 1989; McDowell 1994; Wittgenstein 1921; 1953). But these arguments all depend, in one way or another, upon an anti-realist conception of the mind – claiming, for instance, that since we cannot interpret anyone as entertaining any given fine-grained thought in the absence of linguistic behavior, such thoughts cannot even exist in the absence of such behavior (Davidson 1975). Since the view adopted in this paper – and shared by most cognitive psychologists – is quite strongly realist about thought, I do not propose to devote any time to such arguments.

Notice, too, that Davidson et al. are committed to denying that any non-human animals can entertain genuine thoughts, given that it is very doubtful whether any such animals are capable of understanding and using a natural language (in the relevant sense of "language," that is; see Pinker 1994; Premack 1986). This conclusion conflicts, not just with common sense belief, but also with what can be discovered about animal cognition, both experimentally and by observation of their behavior in the wild (Allen & Bekoff 1997; Byrne 1995; de Waal 1982; 1996; Dickinson & Shanks 1995; Gallistel 1990; Hauser 2000; Povinelli 2000; Savage-Rumbaugh & Lewin 1994; Walker 1983). So not only are the arguments of Davidson et al. unsound, but we have independent reasons to think that their conclusion is false.

Dummett (1994) makes some attempt to accommodate this sort of point by distinguishing between concept-involving thoughts (which are held to be necessarily dependent upon language) and what he calls "proto-thoughts," which are what animals are allowed to possess. Proto-thoughts are said to consist of "visual images superimposed on the visually perceived scene," and are said to be possible only when tied to current circumstances and behavior. But such an account vastly underestimates the cognitive capacities of non-human animals, I believe. If an animal can decide whom to form an alliance with, or can calculate rates of return from different sources of food, or can notice and exploit the ignorance of another, then these things cannot be accounted for in Dummett's terms. And given that conceptual thinking of this sort is possible for animals, then he will be left without any principled distinction between animal thought and human thought.

I do not expect that these brief considerations will convince any of my philosophical opponents, of course; and they aren't meant to. Given the intended readership of this target article, their position is not really one that I need to take seriously. It is mentioned here just to set it aside, and (most important) in order that other, more plausible, versions of the cognitive conception of language shouldn't be confused with it.

I propose, therefore, to take it for granted that thought is conceptually independent of natural language, and that thoughts of many types can actually occur in the absence of such language. But this leaves open the possibility that some types of thought might de facto involve language, given the way in which human cognition is structured. It is on this – weaker but nevertheless still controversial – set of claims that I shall focus. Claims of this type seem to me to have been unjustly under-explored by researchers in the cognitive sciences; partly, no doubt, because they have been run together with the a priori and universalist claims of some philosophers, which have been rightly rejected.

### 3.2. The Joycean machine

Another overly strong form of cognitive conception of language – which has been endorsed by some philosophers and by many social scientists – is that language is, as a matter of fact, the medium of all human conceptual thinking. Most often it has been associated with a radical empiricism about the mind, according to which virtually all human concepts and ways of thinking, and indeed much of the very structure of the human mind itself, are acquired by young children from adults when they learn their native language – these concepts and structures differing widely depending upon the conceptual resources and structures of the natural language in question. This mind-structuring and social-relativist view of language is still dominant in the social sciences, following the writings early in this century of the amateur linguist Whorf (many of whose papers have been collected together in his 1956) – indeed, Pinker (1994) refers to it disparagingly as "the Standard Social Science Model" of the mind.

Perhaps Dennett (1991) is one of the clearest exponents of this view. He argues that human cognitive powers were utterly transformed following the appearance of natural language, as the mind became colonized by *memes* (ideas or concepts, which are transmitted, retained, and selected in a manner supposedly analogous to genes; see Dawkins 1976). Prior to the evolution of language, in this picture, the mind was a bundle of distributed connectionist processors – which conferred on early hominids some degree of flexibility and intelligence, but which were quite limited in their computational powers. The arrival of language then meant that a whole new – serial and compositionally structured – cognitive architecture could be programmed into the system.

This is what Dennett calls the *Joycean machine* (named after James Joyce's "stream of consciousness" writing). The idea is that there is a highest-level processor which runs on a stream of natural-language representations, utilizing learned connections between ideas and patterns of reasoning acquired in and through the acquisition of linguistic memes. According to this account, then, the concept-wielding mind is a kind of social construction, brought into existence through the absorption of memes from the surrounding culture. And in this view, the conceptual mind is both dependent upon, and constitutively involves, natural language.

Admittedly, what Dennett will actually *say* is that animals and pre-linguistic hominids are capable of thought, and engage in much intelligent thinking. But this is because he is not (in my sense) a realist about thoughts. On the contrary, he (like Davidson) is what is sometimes called an "interpretationalist" – he thinks that there is nothing more to thinking than engaging in behavior which is interpretable as thinking. Yet he does seem committed to saying that it is only with the advent of natural language that you get a kind of thinking which involves discrete, structured, semantically-evaluable, causally-effective states – that is, thoughts realistically construed.

Bickerton's proposals (1990; 1995) are somewhat similar, but more biological in flavor. He thinks that, before the evolution of language, hominid cognition was extremely limited in its powers. According to his view, these early forms of hominid cognition consisted largely of a set of relatively simple computational systems, underpinning an array of flexible but essentially behavioristic conditioned responses to stimuli. But then the evolution of language some 100,000 years ago involved a dramatic re-wiring of the hominid brain, giving rise to distinctively human intelligence and conceptual powers.[4]

Bickerton, like Dennett, allows that subsequent to the evolution of language the human mind would have undergone further transformations, as the stock of socially transmitted ideas and concepts changed and increased. But the basic alteration was coincident with, and constituted by, a biological alteration – namely, the appearance of an innately-structured language-faculty. For Bickerton is a nativist about language. (Indeed, his earlier work on the creolization of pidgin languages – 1981 – is often cited as part of an argument for the biological basis of language; see Pinker 1994.) And it is language which, he supposes, conferred on us the capacity for "off-line thinking" – that is, the capacity to think and reason about topics and problems in the abstract, independent of any particular sensory stimulus.

These strong views seem very unlikely to be correct. This is so for two reasons. First, they undervalue the cognitive powers of pre-linguistic children, animals, and earlier forms of hominid. Thus *Homo erectus* and archaic forms of *Homo sapiens,* for example, were able to survive in extremely harsh tundra environments, presumably without language (see below). It is hard to see how this could have been possible without a capacity for quite sophisticated planning and a good deal of complex social interaction (Mithen 1996). Second, the views of Dennett and Bickerton are inconsistent with the sort of central-process modularism which has been gaining increasing support in recent decades. According to this account, the mind contains a variety of conceptual modules – for mind-reading, for doing naïve physics, for reasoning about social contracts, and so on – which are probably of considerable ancestry, pre-dating the appearance of a modular language-faculty.[5] So hominids were already capable of conceptual thought and of reasoning in a complex, and presumably "off-line," fashion *before* the arrival of language.

In sections 3.3 and 3.4 which follow, I shall elaborate briefly on these points. But first, I want to consider a potential reply which might be made by someone sympathetic to Bickerton's position. For Bickerton actually thinks that earlier hominids probably used a form of "proto-language" prior to the evolution of syntax, similar to the language used by young children and to pidgin languages. (This is, in fact, a very plausible intermediate stage in the evolution of natural language.) It might be claimed, then, that insofar as hominids are capable of intelligent thought, this is only because those thoughts are framed in proto-language. So the view that thought is dependent upon language can be preserved.

Such a reply would, indeed, give Bickerton a little extra wiggle-room; but only a little. For as we shall see in section 3.3 below, a good deal of the evidence for hominid thinking is provided by the capacities of our nearest relatives, the great apes, who are known to lack even a proto-language (without a good deal of human enculturation and explicit training, at any rate; Savage-Rumbaugh & Lewin 1994). And some of the other evidence – for example, provided by hominid stone knapping – is not plausibly seen as underpinned by proto-language. Moreover, the various thought-generating central modules, to be discussed in section 3.4 below, are almost surely independent both of language and proto-language. So it remains the case that much hominid thought is independent even of proto-language.

### 3.3. Hominid intelligence

Since social intelligence is something that we share with the other great apes (especially chimpanzees), it is reasonable to conclude that the common ancestor of all apes – and so, by implication, all earlier forms of hominid – will also have excelled in the social domain. While it is still disputed whether chimpanzees have full-blown mind-reading, or "theory of mind," abilities, of the sort attained by a normal four-year-old child, it is not in dispute that the social behavior of great apes can be extremely subtle and sophisticated (Byrne 1995; Byrne & Whiten 1988; 1998; Povinelli 2000).

Two points are worth stressing in this context. One is that it is well-nigh impossible to see how apes can be capable of representing multiple, complex, and constantly changing social relationships (who is friends with whom, who has recently groomed whom, who has recently fallen out with whom, and so on) unless they are capable of structured propositional thought.[6] This is a development of what Horgan and Tienson (1996) call "the tracking argument for Mentalese" (i.e., an argument in support of the claim that thoughts are structured out of recombinable components). Unless the social thoughts of apes were composed out of elements variously representing individuals and their properties and relationships, then it is very hard indeed to see how they could do the sort of one-off learning of which they are manifestly capable. This surely requires separate representations for individuals and their properties and relations, so that the latter can be varied while the former are held constant. So (*contra* Dennett and Bickerton) we have reason to think that all earlier forms of hominid would have been capable of sophisticated conceptual thought (realistically construed), at least in the social domain.[7]

The second point to note is that the social thinking of apes seems sometimes to be genuinely strategic in nature, apparently involving plans that are executed over the course of days or months. Consider, for example, the way in which a band of male chimpanzees will set out quietly and in an organized and purposive manner toward the territory of a neighboring group, apparently with the intention, either of killing some of the males of that group, or of capturing some of its females, or both (Byrne 1995). Or consider the way in which a lower-ranking male will, over the course of a number of months, build up a relationship with the beta male, until the alliance is strong enough for them to cooperate in ousting the alpha male from his position (de Waal 1982). Presumably the thinking that would generate such long-term plans and strategies would have to be "off-line," in the sense of not being tied to or driven by current perceptions of the environment.

We can conclude, then, that all of our hominid ancestors would have had a sophisticated social intelligence. In addition, the stone-tool-making abilities of later species of *Homo erectus* indicate a sophisticated grasp of fracture dynamics and the properties of stone materials. Making stone tools isn't easy. It requires judgment, as well as considerable hand-eye coordination and upper-body strength. And since it uses a reductive technology (starting from a larger stone and reducing it to the required shape) it cannot be routinized in the way that (presumably) nest-building by weaver birds and dam-building by beavers can be. Stone knappers have to hold in mind the desired shape and plan two or more strikes ahead in order to work toward it using variable and unpredictable materials (Mithen 1996; Pelegrin 1993). Moreover, some of the very fine three-dimensional symmetries produced from about half-a-million years ago would almost certainly have required significant capacities for visual imagination – in particular, an ability to mentally rotate an image of the stone product which will result if a particular flake is struck off (Wynn 2000). And this is surely "off-line" thinking if anything is!

We can also conclude that early humans were capable of learning and reasoning about their natural environments with a considerable degree of sophistication. They were able to colonize much of the globe, ranging from Southern Africa to Northwestern Europe to Southeast Asia. And they were able to thrive in a wide variety of habitats (including extremely harsh marginal tundra environments), adapting their life-style to local – and sometimes rapidly changing – circumstances (Mithen 1990; 1996). This again serves as a premise for a version of the "tracking argument," suggesting that early humans were capable of compositionally-structured thoughts about the biological as well as the social worlds.

### 3.4. The modular mind

The above claims about the cognitive powers of our early ancestors both support, and are in turn supported by, the evidence of modular organization in the minds of contemporary humans. According to this account, besides a variety of input and output modules (including, e.g., early vision, face-recognition, and language), the mind also contains a number of innately channeled conceptual modules, designed to process conceptual information concerning particular domains. Although these would not be modules in Fodor's classic (1983) sense, in that they wouldn't have proprietary transducers, might not have dedicated neural hardware, and might not be fully encapsulated, they would still be innately channeled dedicated computational systems, generating information in accordance with algorithms which are not shared with, nor accessible to, other systems.

Plausible candidates for such conceptual modules might include a naïve physics system (Baillargeon 1995; Leslie 1994; Spelke 1994; Spelke et al. 1995), a naïve psychology or "mind-reading" system (Baron-Cohen 1995; Carey 1985; Leslie 1994), a folk-biology system (Atran 1990; 1998; 2002), an intuitive number system (Dehaene 1997; 1995; Gallistel & Gelman 1992; Wynn 1990), a geometrical system for reorienting and navigating in unusual environments (Cheng 1986; Hermer & Spelke 1994; 1996) and a system for processing and keeping track of social contracts (Cosmides & Tooby 1992; Fiddick et al. 2000).

Evidence supporting the existence of at least the first two of these systems (folk-physics and folk-psychology) is now pretty robust. Very young infants already have a set of expectations concerning the behaviors and movements of physical objects, and their understanding of this form of causality develops very rapidly over the first year or two of life. And folk-psychological concepts and expectations also develop very early and follow a characteristic developmental profile. Indeed, recent evidence from the study of twins suggests that three-quarters of the variance in mind-reading abilities amongst three year olds is both genetic in origin and largely independent of the genes responsible for verbal intelligence, with only one quarter of the variance being contributed by the environment (Hughes & Plomin 2000).[8]

Now, *of course* the thesis of conceptual modularity is still highly controversial and disputed by many cognitive scientists. And I cannot pretend to have said enough to have established it here; nor is there the space to attempt to do so. This is going to be one of the large assumptions that I need to ask my readers to take on board as background to what follows. However, there is one sort of objection to conceptual modularity which I should like to respond to briefly here. This is that there simply hasn't been time for all of these modular systems to have evolved (or at any rate, not those of them that are distinctively human – geometry and folk-physics might be the exceptions).

Tomasello (1999) argues that the mere six million years or so since the hominid line diverged from the common ancestor of ourselves and chimpanzees is just too short a time for the processes of evolution to have sculpted a whole suite of conceptual modules. He thinks that explanations of distinctively-human cognition need to postulate just one – at most two – biological adaptations, in terms of which all the other cognitive differences between us and chimpanzees should be explained. His preferred option is theory of mind ability, which underpins processes of cultural learning and cultural accumulation and transmission. Others might argue in similar fashion that the only major biological difference is the language faculty (Perner, personal communication).

The premise of this argument is false, however; six million years is a lot of time, particularly if the selection pressures are powerful ones. (Only 10,000 years separate polar bears and grizzlies, for example.) And this is especially so when, as in the present case, many of the systems in question don't have to be built *ab initio,* but can result from a deepening and strengthening of pre-existing faculties. Thus theory of mind would surely have developed from some pre-existing social-cognition module; folk-biology from a pre-existing foraging system; and so on. In order to reinforce the point, one just has to reflect on the major, and multiple, physical differences between ourselves and chimpanzees – including upright gait, arm length, physical stature, brain size, nasal shape, hairlessness, whites of eyes, and so on. These, too, have all evolved – many of them independently, plainly – over the last six million years.

### 3.5. Taking stock

What has happened in the cognitive sciences in recent decades, then, is this. Many researchers have become increasingly convinced, by neuropsychological and other evidence, that the mind is more or less modular in structure, built up out of isolable, and largely isolated, components (Barkow et al. 1992; Fodor 1983; Gallistel 1990; Hirschfeld & Gelman 1994; Pinker 1997; Sachs 1985; Shallice 1988; Sperber et al. 1995). They have also become convinced that the structure and contents of the mind are substantially innate (Carey 1985; Fodor 1981; 1983; Spelke 1994), and that *language* is one such isolable and largely innate module (Chomsky 1988; Fodor 1983; Pinker 1994). There has then been, amongst cognitive scientists, a near-universal reaction against the cognitive conception of language, by running it together with the Whorfian hypothesis. Most researchers have assumed, without argument, that if they were to accept any form of cognitive conception of language, then that would commit them to Whorfian linguistic relativism and radical empiricism, and would hence be inconsistent with their well-founded beliefs in modularity and nativism (Pinker 1994).

It is important to see, however, that someone endorsing the cognitive conception of language does not *have* to regard language and the mind as cultural constructs, either socially determined or culturally relative. In fact, some form of cognitive conception of language can equally well be deployed along with a modularist and nativist view of language and mind. There is a range of positions intermediate between the input-output conception of language on the one hand and Whorfian relativism (the Standard Social Science Model) on the other, which deserve the attention of philosophers and cognitive scientists alike. These views are nativist as opposed to empiricist about language and much of the structure of the mind, but nevertheless hold that language is constitutively employed in many of our thoughts.

## 4. Language and conscious thinking

What is at stake, then, is the question whether language might be constitutively involved in *some* forms of human thinking. But which forms? In previous work I suggested that language might be the medium in which we conduct our *conscious propositional* thinking – claiming, that is, that inner speech might be the vehicle of conscious-conceptual (as opposed to conscious visuo-spatial) thinking (Carruthers 1996). This view takes seriously and literally the bit of folk-wisdom with which this paper began – namely, that much of our conscious thinking (namely, our propositional thinking) is conducted in inner speech.

Now, if the thesis here is that the cognitive role of language is *confined* to conscious thinking, then it will have to be allowed that much propositional thinking also takes place independently of natural language – for it would hardly be very plausible to maintain that there is no thinking but conscious thinking. And there are then two significant options regarding the relations between non-conscious language-independent thought, on the one hand, and conscious language-involving thinking, on the other. For either we would have to say that anything which we can think consciously, in language, can also be thought non-consciously, independently of language; or we would have to say that there are some thought-types which can only be entertained at all, by us, when tokened consciously in the form of an imaged natural language sentence.

Suppose that it is the first – weaker and more plausible – of these options that is taken. Then we had better be able to identify some element of the distinctive causal role of an imaged sentence that is sufficiently thought-like or inference-like for us to be able to say that the sentence in question is partly *constitutive of* the (conscious) tokening of the thought-type in question, rather than being merely *expressive of* it. For otherwise – if everything which we can think consciously, in language, we can also think non-consciously, without language – what is to block the conclusion that inner speech is merely the means by which we have access to our occurrent thoughts, without inner speech being in any sense constitutive of our thinking? (On this, at length, see Carruthers 1998b.)

There would seem to be just two distinct (albeit mutually consistent) possibilities here. One (implicit in Carruthers 1996) would be to propose a suitably weakened version of Dennett's *Joycean machine* hypothesis. While allowing (*contra* Dennett) that much conceptual thinking (realistically construed) and all conceptual thought-types are independent of language (in the sense of not being constituted by it), we could claim that there are certain learned habits and patterns of thinking and reasoning that are acquired linguistically, and then restricted to linguistic (and conscious) tokenings of the thoughts that they govern. It is surely plausible, for example, that exact long-division or multiplication can only be conducted consciously, in imaged manipulations of numerical symbols. Similarly, it may be that the result of taking a course in logic is that one becomes disposed to make transitions between sentences, consciously in language, where one would otherwise not have been disposed to make the corresponding transitions between the thoughts expressed. If these sorts of possibilities are realized, then we would have good reason to say of a token application of a particular inference-form, that the imaged natural language sentences involved are constitutive of the inference in question, since it could not have taken place without them.

A second possibility is proposed and defended by Frankish (1998a; 1998b; forthcoming; see also Cohen 1993). This is that the distinctive causal role of inner speech is partly a function of our *decisions* to accept, reject, or act on the propositions which our imaged sentences express. I can frame a hitherto unconsidered proposition in inner speech and decide that it is worthy of acceptance, thereby committing myself to thinking and acting thereafter as if that sentence were true. Then provided that I remember my commitments and execute them, it will be *just as if* I believed the proposition in question. (In his published work, Frankish describes this level of mentality as the "virtual mind" and the beliefs in question as "virtual beliefs.") But, by hypothesis, I would never have come to believe what I do, nor to reason as I do reason, except via the tokening of sentences in inner speech. Frankish argues, in effect, that there is a whole level of mentality (which he now dubs "supermind") which is constituted by our higher-order decisions and commitments to accept or reject propositions; and that language is constitutive of the thoughts and beliefs which we entertain at this level.

Such views have considerable plausibility; and it may well be that one, or other, or both of these accounts of the causal role of inner speech is correct. Indeed, the *dual process theory* of human reasoning developed over the years by Evans and colleagues (Evans & Over 1996; Wason & Evans 1975), and more recently by Stanovich (1999), combines elements of each of them. According to this account, in addition to a suite of computationally powerful, fast, and implicit reasoning systems (from our perspective, a set of conceptual modules), the mind also contains a slow, serial, and explicit reasoning capacity, whose operations are conscious and under personal control, and which is said (by some theorists at least; e.g., Evans & Over 1996) to involve natural language. The emphasis here on learned rules in the operations of the explicit system is reminiscent of Dennett's "Joycean machine," whereas the stress on our having personal control over the operations of that system seems very similar to Frankish's conception of "supermind."

Not only is some form of dual-process theory plausible, but it should also be stressed that these accounts are independent of central-process modularism. Those who deny the existence of any conceptual modules can still accept that there is a level of thinking and reasoning which is both language-involving and conscious. It is surely plain, however, that none of the above accounts can amount to the most fundamental cognitive function of language once conceptual modularity is assumed.

Given conceptual modularity, then unless the above views are held together with the thesis to be developed in section 5 below – namely, that language provides the medium for inter-modular communication and non-domain-specific thinking – then we can set their proponents a dilemma. Either they must claim that a domain-general architecture was in place prior to the evolution of language; or they must allow that there was no significant domain-general cognition amongst hominids prior to the appearance of language and language-involving conscious thinking; and they must claim that such cognition still evolved as a distinct development, either at the same time or later. Since contemporary humans are manifestly capable of conjoining information across different domains in both their theoretical thought and their planning, then either pre-linguistic humans must already have had domain-general theoretical and practical reasoning faculties; or they must have evolved them separately at the same time or after the evolution of the language faculty (that is, if it isn't language itself which enables us to combine information across modules).

The problem with the first alternative, however – namely, that domain-general reasoning capacities pre-dated language – is that the evidence from cognitive archaeology suggests that this was not the case. For although the various sub-species of *Homo erectus* and archaic forms of *Homo sapiens* were smart, they were not *that* smart. Let me briefly elaborate.

As Mithen (1996) demonstrates at length and in detail, the evidence from archaeology is that the minds of early humans were in important respects quite unlike our own. While they successfully colonized diverse and rapidly changing environments, the evidence suggests that they were incapable of bringing together information across different cognitive domains. It seems that they could not (or did not) mix information from the biological world (utilized in hunting and gathering) with information about the physical world (used in tool making); and that neither of these sorts of information interacted with their social intelligence. Although they made sophisticated stone tools, they did not use those tools for specialized purposes (with different kinds of arrow-head being used for different kinds of game, for example); and they did not make tools out of animal products such as antler and bone. There is no sign of the use of artifacts as social signals, in the form of body ornaments and such, which is ubiquitous in modern human cultures. And there is no indication of totemization or other sorts of linkages between social and animal domains, such as lion-man figurines, cave-paintings, or the burying of the dead with (presumably symbolic) animal parts – which all emerge onto the scene for the first time with modern humans. As Mithen summarizes the evidence, it would appear that early humans had sophisticated special intelligences, but that these faculties remained largely isolated from one another.

The problem with the second horn of the dilemma sketched above is just that it is hard to believe, either that a domain-general reasoning faculty might have evolved *after* the appearance of language some 100,000 years ago (in just the 20,000 years or so before the beginning of the dispersal of modern humans around the globe), or that language and domain-general capacities might have co-evolved as distinct faculties. For as we shall see in section 5, the evolution of language would in any case have involved the language faculty taking inputs from, and sending outputs to, the various modular systems, if there wasn't already a domain-general system for it to be linked to. And it is hard to discern what the separate selection pressures might have been, which would have led to the development of two distinct faculties at about the same time (language and domain-general thought), when just one would serve.

## 5. Language as the medium of non-domain-specific thinking

The hypothesis that I particularly want to explore, then, is that natural language is the medium of non-domain-specific thought and inference. Versions of this hypothesis have been previously proposed by Carruthers (1996; 1998a), by Mithen (1996), and by Spelke and colleagues (Hermer-Vazquez et al., 1999; Spelke & Tsivkin 2001; Spelke, in press). I shall sketch the thesis itself, outline the existing experimental evidence in its support, and then (in the section following) consider some of its ramifications and possible elaborations. Finally (in sect. 7), I shall discuss what further evidence needs to be sought as a test of our thesis.

### 5.1. The thesis

The hypothesis in question assumes a form of central-process modularism. That is, it assumes that in addition to the various input and output modules (vision, face-recognition, hearing, language, systems for motor-control, etc.), the mind also contains a range of conceptual modules, which take conceptual inputs and deliver conceptual outputs. Evidence of various sorts has been accumulating in support of central-process modularism in recent decades (some of which has already been noted above). One line of support is provided by evolutionary psychologists, who have argued on both theoretical and empirical grounds that the mind contains a suite of domain-specific cognitive adaptations (Barkow et al., 1992; Pinker 1997; Sperber 1996;). But many who would not describe themselves as "evolutionary psychologists" have argued for a modular organization of central cognition, on developmental, psychological, and/or neuro-pathological grounds (Baron-Cohen 1995; Carey 1985; Carey & Spelke 1994; Gallistel 1990; Hauser & Carey 1998; Leslie 1994; Smith & Tsimpli 1995; Shallice 1988; Spelke 1994).

What cognitive resources were antecedently available, then, prior to the evolution of the language faculty? Taking the ubiquitous laboratory rat as a representative example, I shall assume that all mammals, at least, are capable of thought – in the sense that they engage in computations which deliver structured (propositional) belief-like states and desire-like states (Dickinson 1994; Dickinson & Balleine 2000). I shall also assume that these computations are largely carried out within modular systems of one sort or another (Gallistel 1990) – after all, if the project here is to show how cross-modular thinking in humans can emerge out of modular components, then we had better assume that the initial starting-state was a modular one. Furthermore, I shall assume that mammals possess some sort of simple non-domain-specific practical reasoning system, which can take beliefs and desires as input, and figure out what to do.

I shall assume that the practical reasoning system in animals (and perhaps also in us) is a relatively simple and limited-channel one. Perhaps it receives as input the currently-strongest desire and searches amongst the outputs of the various belief-generating modules for something which can be done in relation to the perceived environment which will satisfy that desire. So its inputs have the form DESIRE [Y] and BELIEF [IF X THEN Y], where X should be something for which an existing motor-program exists. I assume that the practical reasoning system is *not* capable of engaging in other forms of inference (generating new beliefs from old), nor of combining together beliefs from different modules; though perhaps it *is* capable of chaining together conditionals to generate a simple plan – for example, BELIEF [IF W THEN X], BELIEF [IF X THEN Y], → BELIEF [IF W THEN Y].

The central modules will take inputs from perception, of course. And my guess is that many of the beliefs and desires generated by the central modules will have partially indexical contents – thus a desire produced as output by the sex module might have the form, "I want to mate with *that* female," and a belief produced by the causal-reasoning module might have the form, "*That* caused *that*." So if the practical reasoning system is to be able to do anything with such contents, then it, too, would need to have access to the outputs of perception, to provide anchoring for the various indexicals. The outputs of the practical reasoning system are often likely to be indexical too, such as an intention of the form, "I'll go *that* way."

The inputs to central-process modules can presumably include not only conceptualized perceptions but also propositional descriptions (in the latter case deriving from linguistic input – for we surely use our mind-reading system, for example, when processing a description of someone's state of mind as well as when observing their behavior). And in some cases, too, the inputs to a module will include the outputs of other central-process modules; for we might expect that there will be cases in which modules are organized into some sort of hierarchy. But what of the *outputs* from central-process modules? Besides being directed to other modules (in some instances), and also to the practical reasoning system, where is the information which is generated by central-process modules normally sent? And in particular, is there some *non*-domain-specific central arena where all such information is collated and processed?

The hypothesis being proposed here is that there *is* such an arena, but one which crucially implicates natural language, and which cannot operate in the absence of such language. Moreover, the hypothesis is not just that our conscious propositional thinking involves language (as sketched in sect. 4 above), but that *all* non-domain-specific reasoning of a non-practical sort (whether conscious or non-conscious) is conducted in language. And as for the question of what a non-conscious tokening of a natural language sentence would be like, we can propose that it would be a representation stripped of all imagistic-phonological features, but still consisting of natural language lexical items and syntactic structures. (The role of syntax in the present account will be further explored in sect. 6.1 below.)

Chomsky (1995) has maintained, for example, that there is a level of linguistic representation which he calls "logical form" (LF), which is where the language faculty interfaces with central cognitive systems. We can then claim that all cross-modular thinking consists in the formation and manipulation of these LF representations. The hypothesis can be that all such thinking operates by accessing and manipulating the representations of the language faculty. Where these representations are *only* in LF, the thoughts in question will be non-conscious ones. But where the LF representation is used to generate a full-blown phonological representation (an imaged sentence), the thought will nomally be conscious. And crucially for my purposes, the hypothesis is that the language faculty has access to the outputs of the various central-process modules, in such a way that it can build LF representations which combine information across domains.

Let me say a just little more about the conscious/non-conscious distinction as it operates here. As I shall mention again in a moment (and as I shall return to at some length in sect. 6.2) language is both an input and an output module. Its production sub-system must be capable of receiving outputs from the conceptual modules in order to transform their creations into speech. And its comprehension sub-system must be capable of transforming heard speech into a format suitable for processing by those same conceptual modules. Now when LF representations built by the production sub-system are used to generate a phonological representation, in "inner speech," that representation will be consumed by the comprehension sub-system, and made available to central systems. One of these systems is a theory of mind module. And on the sort of higher-order theory of consciousness which I favor (Carruthers 2000), perceptual and imagistic states get to be phenomenally conscious by virtue of their availability to the higher-order thoughts generated by the theory of mind system (i.e., thoughts *about* those perceptual and imagistic states). So this is why inner speech of this sort is conscious: It is because it is available to higher-order thought.

The hypothesis, then, is that non-domain-specific, cross-modular, propositional thought depends upon natural language – and not just in the sense that language is a necessary condition for us to entertain such thoughts, but in the stronger sense that natural language representations are the bearers of those propositional thought-contents. So language is constitutively involved in (some kinds of) human thinking. Specifically, language is the vehicle of non-modular, non-domain-specific, conceptual thinking which integrates the results of modular thinking.

Before moving on to discuss the evidence in support of our thesis, consider one further question. Why does it have to be language, and not, for example, visual imagery that serves the integrative function? For visual images, too, can carry contents which cross modular domains. But such visual thinking will access and deploy the resources of a peripheral input module. It cannot, therefore, play a role in integrating information across conceptual modules, because the latter exist down-stream of the input-systems. Vision provides input to conceptual modules, and doesn't receive output from them. The language faculty, in contrast, while also "peripheral," has both input and output functions. (I shall return to this point again in sect. 6.2 below.) I would hypothesize, therefore, that in cases where visual images have cross-modular contents (and aren't memory images), they are always generated from some linguistic representation which originally served to integrate those contents.

### 5.2. The evidence

What evidence is there to support the hypothesis that natural language is the medium of inter-modular communication, or of non-domain-specific integrated thinking? Until recently, the evidence was mostly circumstantial. For example, one indirect line of argument in support of our thesis derives from cognitive archaeology, when combined with the evidence of contemporary central-process modularism (Mithen 1996). For as we noted above, it seems that we have significant evidence of cross-modular thought only following the emergence of contemporary humans some 100,000 years ago; whereas independent evidence suggests that language, too, was a late evolutionary adaptation, only finally emerging at about the same time (perhaps from an earlier stage of "proto-language" – Bickerton 1990; 1995). So the simplest hypothesis is that it is language which actually enables cross-modular thinking.

Another strand of indirect evidence can be provided if we take seriously the idea that the stream of inner verbalization is constitutive of (some forms of) thinking (Carruthers 1996). For as we saw in section 4, such views can only plausibly be held (given the truth of central-process modularism) together with the present hypothesis that language is the main medium of inter-modular communication.

Much more important, however, direct tests of (limited forms of) our hypothesis have now begun to be conducted. The most important of these is Hermer-Vazquez et al. (1999), which provides strong evidence that the integration of geometric properties with other sorts of information (color, smell, patterning, etc.) is dependent upon natural language. The background to their studies with human adults is the apparent discovery of a *geometric module* in rats by Cheng (1986), as well as the discovery of a similar system in pre-linguistic human children (Hermer & Spelke 1994; 1996).

Cheng (1986) placed rats in a rectagonal chamber, and allowed them to discover the location of a food source. They were then removed from the chamber and disoriented, before being placed back into the box with the food now hidden. In each case there were multiple cues available – both geometric and non-geometric – to guide the rats in their search. For example, the different walls might be distinctively colored or patterned, one corner might be heavily scented, and so on. In fact in these circumstances the rats relied exclusively on geometric information, searching with equal frequency, for example, in the two geometrically-equivalent corners having a long wall on the left and a short wall on the right. Yet rats are perfectly well capable of noticing and remembering non-geometric properties of the environment and using them to solve other tasks. So it appears that, not only are they incapable of integrating geometric with non-geometric information in these circumstances, but that geometric information takes priority.

(This makes perfectly good ecological-evolutionary sense. For in the rat's natural environment, overall geometrical symmetries in the landscape are extremely rare, and geometrical properties generally change only slowly with time; whereas object-properties of color, scent-markings, and so on will change with the weather and seasons. So a strong preference to orient by geometrical properties is just what one might predict.)

Hermer and Spelke (1994; 1996) found exactly the same phenomenon in pre-linguistic human children. Young children, too, rely exclusively on geometric information when disoriented in a rectangular room, and appear incapable of integrating geometrical with non-geometrical properties when searching for a previously seen but now-hidden object. Older children and adults are able to solve these problems without difficulty – for example, they go straight to the corner formed with a long wall to the left and a short *blue* wall to the right. It turns out that success in these tasks isn't directly correlated with age, nonverbal IQ, verbal working-memory capacity, vocabulary size, or comprehension of spatial vocabulary. In contrast, the only significant predictor of success in these tasks which could be discovered, was spontaneous use of spatial vocabulary conjoined with object-properties (e.g., "It's left of the red one"). Even by themselves, these data strongly suggest that it is language which enables older children and adults to integrate geometric with non-geometric information into a single thought or memory.

Hermer-Vazquez et al. (1999) set out to test this idea with a series of dual-task experiments with adults. In one condition, subjects were required to solve one of these orientation problems while shadowing (i.e., repeating back) speech played to them through a set of headphones. In another condition, they were set the same problems while shadowing (with their hands) a *rhythm* played to them in their headphones. The hypothesis was that speech-shadowing would tie up the resources of the language faculty, whereas the rhythm-shadowing tasks would not; and great care was taken to ensure that the latter tasks were equally if not more demanding of the resources of working memory.

The results of these experiments were striking. Shadowing of speech severely disrupted subjects' capacity to solve tasks requiring integration of geometric with non-geometric properties. In contrast, shadowing of rhythm disrupted subjects' performance relatively little. Moreover, a follow-up experiment demonstrated that shadowing of speech didn't disrupt subjects' capacities to utilize non-geometric information *per se* – they were easily able to solve tasks requiring only memory for object-properties. So it would appear that it is language itself which enables subjects to conjoin geometric with non-geometric properties, just as the hypothesis that language is the medium of cross-modular thinking predicts.

Of course, this is just one set of experiments – albeit elegant and powerful – concerning the role of language in enabling information to be combined across just two domains (geometrical and object-properties). In which case, little direct support is provided for the more-demanding thesis that language serves as the vehicle of inter-modular integration in general. But the evidence does at least suggest that the more general thesis may be well worth pursing.

### 5.3. Challenging the data

The position taken by Hermer-Vazquez et al. (1999) has come under pressure from two different directions. First, there are claims that other species (chickens, monkeys) can integrate geometric and landmark information when disoriented (Gouteux et al. 2002; Vallortigara et al. 1990). And second, there is the finding that success in these tasks amongst young children is somewhat sensitive to the size of the room – in a larger room, significantly more young (four-year-old) children make the correct choice, utilizing both geometric and landmark information; and even more five- and six-year-old children are also able to make the correct choice (Learmonth et al. 2001; in press).

To begin unpicking the significance of these new results, we need to return to some of the original claims. It is too strong to say that the original data with rats (Cheng, 1986) showed the existence of a geometric module in that species. For rats can use landmark information when navigating in other circumstances. The fact is just that they don't use such information when disoriented. Nor is it established that rats *cannot* integrate geometric with landmark information. The fact is just that they *do not* utilize both forms of information when disoriented. So the data are consistent with the following model: there are no modules; rather, geometric and landmark information are both processed according to general-purpose algorithms and made available to some sort of practical reasoning system. But when disoriented, rats only pay attention to, and only make use of, the geometric information.

Even if one thinks (as I do) that other forms of evidence and other arguments make some sort of modularist architecture quite likely, the following proposal is still consistent with the data: both the geometric and landmark modules normally make their information available to some sort of practical reasoning system; but when disoriented, rats show a strong preference to make use *only* of the geometric information.

Equally, however, the fact that other species are able to solve these problems doesn't show that members of those species can integrate geometric with landmark information into a single belief or thought. For it is possible to solve these tasks by making use of the information sequentially. The problems can be solved by first re-orienting to the landmark, and then using geometric information to isolate the correct corner. So the data are consistent with the following modularist model: Both the geometric and landmark modules make their outputs available to a limited-channel practical reasoning system, where the latter doesn't have the inferential resources to integrate information from different modules; rather, it can only utilize that information sequentially, using a variety of heuristics (both innate and learned) in selecting the information to be used, and in what order. According to this view, the difference between monkeys and rats is just that the former utilize landmark information first, before using geometry; whereas the latter use geometry exclusively in these circumstances. Neither species may in fact be capable of integrating geometrical with landmark information.

(It is tempting to seek an adaptionist explanation of these species differences. Open-country dwellers such as rats and pre-linguistic humans may have an innate pre-disposition to rely only on geometric information when disoriented because such information alone will almost always provide a unique solution – given that rectagonal rooms don't normally occur in nature! Forest dwellers such as chickens and monkeys, in contrast, have an innate pre-disposition to seek for landmark information first, only using geometric information to navigate in relation to a known landmark. This is because geometric information is of limited usefulness in a forest – the geometry is just too complex to be useful in individuating a place in the absence of a landmark such as a well-known fruit-tree.)

What of the new data concerning the effects of room size? Well, the first thing to say is that this data leaves intact the finding by Hermer-Vazquez et al. (1999) that the best predictor of success in children (in small room experiments) is productive use of left-right vocabulary. This suggests, both that language has something to do with their success, and that it is specifically syntax (the capacity to integrate different content-bearing items into a single thought) that is required. For if the role of language were simply to help fix the salience and importance of landmark information, one would expect that it should have been productive use of color vocabulary, rather than spatial vocabulary, which was the best predictor of success. For by hypothesis, after all, children are already disposed to use geometric information in reorienting; their problem is to make use of color information as well.

Equally untouched are the experiments with adults involving speech shadowing and rhythm shadowing, which found that the former greatly disrupts the capacity to use geometric and landmark information together, whereas the latter does not. These results, too, suggest that it is language which enables adults to integrate the two forms of information.

Why should room size have any effect upon children's performance, however? Here is one testable possibility, which is consistent with the theoretical framework of Hermer-Vazquez et al. (1999) and the present author. In a small room (4 feet by 6 feet) it requires but very little time and energy to select a corner and turn over a card. In a larger room (8 feet by 12 feet), in contrast, children have to take a few steps in order to reach a selected corner, giving them both a motive, and the time, to reflect. It may then be that the children who were able to succeed in the large-room condition were on the cusp of having the linguistic competence necessary to integrate geometric and landmark information. Perhaps they could do this, but only haltingly and with some effort. Then it is only to be expected that such children should succeed when given both the time and the motive to do so.

### 5.4. More data: Language and arithmetic

We have examined one set of data which provides strong support for a limited version of our thesis. Data is now available in one other domain – that of number. This comes from a recent bilingual training study conducted by Spelke and Tsivkin (2001). The background to this study is the discovery of numerical capacities in animals and human infants, of two different sorts. One is the capacity possessed by many different kinds of animal (including birds and fish) to represent the approximate numerosity of largish sets of items (Dehaene 1997; Gallistel 1990). This capacity is utilized especially in foraging, enabling animals to estimate rates of return from different food sources. The other numerical capacity is possessed by monkeys and human infants, at least. It is a capacity to represent the exact number of small sets of items (up to about four), keeping track of their number using simple forms of addition and subtraction (Gallistel & Gelman 1992; Hauser & Carey 1998).

The developmental hypothesis which forms the backdrop to Spelke and Tsivkin's study is that language-learning in human children – specifically, learning to pair number words with items in a set through the process of counting – builds upon these two pre-linguistic numerical capacities to enable humans to represent exact numbers of unlimited magnitude. But the developmental hypothesis could be interpreted in two ways. On one interpretation, the role of language is to load the child's mind with a set of language-independent exact numerical concepts – so that, once acquired, the capacity to represent exact large magnitudes is independent of language. The other interpretation is that it is the numerical vocabulary of a specific natural language which forms the medium of exact-magnitude representation, in such a way that natural language is the vehicle of arithmetic thought. It is this latter interpretation which Spelke and Tsivkin (2001) set out to test.

They conducted three different bilingual arithmetic training experiments. In one experiment, bilingual Russian-English college students were taught new numerical operations; in another, they were taught new arithmetic equations; and in the third, they were taught new geographical and historical facts involving both numerical and non-numerical information. After learning a set of items in each of their two languages, subjects were tested for knowledge of those and of new items in both languages. In all three studies subjects retrieved information about exact numbers more effectively in the language in which they were trained on that information, and they solved trained problems more effectively than new ones. In contrast, subjects retrieved information about approximate numbers and about non-numerical (geographical or historical) facts with equal ease in their two languages, and their training on approximate number facts generalized to new facts of the same type.[9] These results suggest that one or another natural language is the vehicle of thought about exact numbers, but not for representing approximate numerosity (a capacity shared with other animals).

What we have, then, is the beginnings of evidence in support of our general thesis that natural language is the medium of inter-modular non-domain-specific thinking. In sect. 7, I shall briefly consider where and how one might search for yet more evidence. But first I shall look at some of the questions raised by our thesis.

## 6. Ramifications and implications

In this section I shall consider – very briefly – some of the implications of the hypothesis just proposed, as well as discussing a number of outstanding questions.

### *6.1. Speech production*

It is plain that the present hypothesis commits us to a non-classical account of speech production. Classically, speech begins with thought – with a mental representation of the message to be communicated; and then linguistic resources (lexical and phonological items, syntactic structures and so on) are recruited in such a way as to express that thought in speech. (See, e.g., Levelt 1989.) But of course this is a picture which we cannot endorse. We cannot accept that the production of the sentence, "The toy is to the left of the blue wall" begins with a tokening of the thought, THE TOY IS TO THE LEFT OF THE BLUE WALL (in Mentalese), since our hypothesis is that such a thought cannot be entertained independently of being framed in natural language.[10] How, then, does the sentence get assembled? I have to confess that I don't have a complete answer to this question in my pocket at the moment! But then this need be no particular embarrassment, since classical theorists don't have an account of how *their* initial Mentalese thoughts are assembled, either.

In fact our hypothesis enables us to split the problem of speech production/domain-general thought-generation into two, each of which may prove individually more tractable. For we are supposing that central modules are capable of generating thoughts with respect to items in their domain. Thus the geometrical module might build a thought of the form, THE TOY IS IN THE CORNER WITH A LONG WALL ON THE LEFT AND A SHORT WALL ON THE RIGHT, whereas the object-property system might build a thought of the form, THE TOY IS BY THE BLUE WALL. Each of these thoughts can be taken as input by the language faculty, we may suppose, for direct translation into natural language expression. It is, then, not so *very* difficult to suppose that the language faculty might have the resources to combine these two thoughts into one, forming a representation with the content, *The toy is in the corner with a long wall on the left and a short blue wall on the right.*[11] Nor is this such a *very* large departure from classical accounts. (Certainly it is much less radical than Dennett's endorsement of *pandemonium models* of speech production; see his 1991.)

The two tasks facing us, then, in explaining speech production are, first, to explain how the thoughts generated by central modules are used to produce a natural language sentence with the same content; and second, to explain how the language faculty can take two distinct sentences, generated from the outputs of distinct conceptual modules, and combine them into a single natural language sentence. There is some reason to hope that, thus divided, the problem may ultimately prove tractable. In responding to the first part of the problem we *can* utilize classical accounts of speech production. Here just let me say a brief word about the second of the above problems, by way of further explaining the role of syntax in my model.

Two points are suggestive of how distinct domain-specific sentences might be combined into a single domain-general one. One is that natural language syntax allows for multiple embedding of adjectives and phrases. Thus one can have, "The food is in the corner with the *long* wall on the left," "The food is in the corner with the *long straight* wall on the left," and so on. So there are already "slots" into which additional adjectives – such as "blue" – can be inserted. The second point is that the reference of terms like "the wall," "the food," and so on will need to be secured by some sort of indexing to the contents of current perception or recent memory. In which case it looks like it would not be too complex a matter for the language production system to take two sentences sharing a number of references like this, and combine them into one sentence by inserting adjectives from one into open adjective-slots in the other. The language faculty just has to take the two sentences, "The food is in the corner with the long wall on the left" and, "The food is by the blue wall" and use them to generate the sentence, "The food is in the corner with the long *blue* wall on the left," or the sentence, "The food is in the corner with the long wall on the left *by the blue wall.*"

### *6.2. Cycles of LF activity*

The present proposals may enable us to rescue one other aspect of Dennett's (1991) "Joycean machine" hypothesis (in addition to the "learned linguistic habits" idea, discussed in sect. 4 above), again without commitment to his claim that language is the medium of *all* (realistically-construed and structured) conceptual thought. This is the suggestion that by asking ourselves questions we can initiate searches in a number of different modular systems, perhaps generating new information which in turn generates new questions, and so on. In general this cycle of questions and answers will go on consciously, in inner speech (as Dennett supposes), but it might also be conducted non-consciously (either below the threshold of attention or *perhaps* in LF – I shall not pursue this suggestion here), through over-learning. Let me elaborate.

The crucial point for these purposes is that natural language is *both* an input *and* an output system. It is the output sub-system of the language faculty which will initially play the role of conjoining information from different conceptual modules, since it is this sub-system which will have been designed to receive inputs from those modules. This is because the evolution of a language system would already have required some sort of interface between the Mentalese outputs of the conceptual modules and the speech-production sub-system of the language faculty, so that those thoughts could receive expression in speech. And we are supposing that this interface became modified during the evolution of language so that thoughts deriving from distinct conceptual modules could be combined into a single natural language sentence. But when the resulting LF representation is used to generate a phonological representation of that sentence, in inner speech, this might normally co-opt the resources of the *input* sub-system of the language faculty, in such a way as to generate a "heard" sentence in auditory imagination. By virtue of being "heard," then, the sentence would also be taken as *input* to the conceptual modules which are down-stream of the comprehension sub-system of the language faculty, receiving the latter's output. So the cycle goes: thoughts generated by central modules are used to frame a natural language representation, which is used to generate a sentence in auditory imagination, which is then taken as input by the central modules once again.

A comparison with visual imagination may be of some help here. According to Kosslyn (1994), visual imagination exploits the top-down neural pathways (which are deployed in normal vision to direct visual search and to enhance object recognition) in order to generate visual stimuli in the occipital cortex, which are then processed by the visual system in the normal way, just as if they were visual percepts. A conceptual or other non-visual representation (of the letter "A," as it might be) is projected back through the visual system in such a way as to generate activity in the occipital cortex, just as if a letter "A" were being perceived. This activity is then processed by the visual system to yield a quasi-visual percept.

Something very similar to this presumably takes place in auditory (and other forms of) imagination. Back-projecting neural pathways which are normally exploited in the processing of heard speech will be recruited to generate a quasi-auditory input, yielding the phenomenon of "inner speech." In this way the *outputs* of the various conceptual modules, united into a sentence of LF by the production sub-system of the language faculty, can become *inputs* to those same modules by recruiting the resources of the comprehension sub-system of the language faculty, in inner speech.

But now, how would a sentence which combines information across a number of distinct central-modular domains have its content "split up" so as to be taken as input again by those modules, with their proprietary and domain-specific concepts? One plausible suggestion is that this is one of the primary functions of so-called "mental models" – non-sentential, quasi-imagistic representations of the salient features of a situation being thought about (Johnson-Laird 1983). For it is now well-established that mental models play an indispensable role in discourse comprehension (see Harris 2000, for reviews). When listening to speech, what people do is construct a mental model of the situation being described, which they can then use to underpin further inferences. The reason why this may work is that mental models, being perception-like, are already of the right *form* to be taken as input by the suite of conceptual modules. For of course those modules would originally have been built to handle perceptual inputs, prior to the evolution of language.

So the suggestion is that language, by virtue of its role in unifying the outputs of conceptual modules, and by virtue of our capacity for auditory imagination, can be used to generate cycles of central-modular activity, hence recruiting the resources of a range of specialized central-modular systems in seeking solutions to problems. This may be one of the main sources of the cognitive flexibility and adaptability which is so distinctive of our species. But how, exactly, are the LF questions which are used in such cycles of enquiry to be generated? How does the language system formulate interrogative sentences which are both relevant and fruitful? I do not have an answer to this question. But I am not embarrassed by this lack, since I suspect that no one has, as yet, a worked-out story about how interrogative thoughts are formed.

### 6.3. LF consumers

What are the consumer systems for the LF sentences generated from the outputs of the central modules? What can be done with an LF sentence, once it has been formulated? One thing which it can be used for, obviously, is to generate an imaged natural language sentence with the same content, thereby rendering the thought in question conscious, and triggering the kinds of mental activity and further consequences distinctive of conscious thinking. Specifically, it may make possible sequences of thought in accordance with learned habits or rules (see sect. 4 above); it will make that sentence and its content accessible to a variety of central-process systems for consideration, and for acceptance or rejection (see sect. 4 above); and it may make possible cycles of LF activity involving central modules, generating new thought-contents which were not previously available (see sect. 6.2 above).

In addition, I can think of two plausible special-purpose systems which may have been designed to consume LF representations. First, it may be that there is a domain-general factual memory system. (It is already known that there is a factual-semantic memory system which is distinct from the experiential-personal memory system, and that the latter is experience-driven whereas the former is not. See Baddeley 1988.) This system would either store domain-general information in the form of LF sentences, or (more plausibly) in some other format (mental models?) generated by the LF sentences which it takes as input. (Recall the data from

Spelke & Tsivkin 2001, that geographical and historical information is recalled equally readily whether or not the language of learning is the same as the language of testing.)

Second, it may be that there is, in addition, some sort of innately-channeled abductive reasoning faculty, which places constraints upon sentence acceptance (Carruthers 1992; 2002). This would be a domain-general reasoning system, taking LF sentences as input and generating LF sentences as output. The reasons for believing in a faculty of "inference to the best explanation" are two-fold. First, there are certain very general constraints on theory-choice employed in science which are equally valid in other areas of enquiry, and which appear to be universal amongst humans. While no one any longer thinks that it is possible to codify these principles of abductive inference, it is generally agreed that the good-making features of a theory include such features as: (1) *accuracy* (predicting all or most of the data to be explained, and explaining away the rest); (2) *simplicity* (being expressible as economically as possible, with the fewest commitments to distinct kinds of fact and process); (3) *consistency* (internal to the theory or model); (4) *coherence* (with surrounding beliefs and theories, meshing together with those surroundings, or at least being consistent with them); (5) *fruitfulness* (making new predictions and suggesting new lines of enquiry); and (6) *explanatory scope* (unifying together a diverse range of data).

Essentially the same principles are employed in many contexts of everyday reasoning. Most strikingly for our purposes, however, such principles are employed by human hunter-gatherers, especially when tracking prey. Successful hunters will often need to develop speculative hypotheses concerning the likely causes of the few signs available to them, and concerning the likely future behavior of the animal; and these hypotheses are subjected to extensive debate and further empirical testing by the hunters concerned. When examined in detail these activities look a great deal like science, as Liebenberg (1990) demonstrates.

First, there is the invention of one or more hypotheses concerning the unobserved (and now unobservable) causes of the observed signs, and the circumstances in which they may have been made. These hypotheses are then examined and discussed for their accuracy, coherence with background knowledge, and explanatory and predictive power.[12] One of them may emerge out of this debate as the most plausible, and this can then be acted upon by the hunters, while at the same time searching for further signs which might confirm or count against it. In the course of a single hunt one can then see the birth, development, and death of a number of different "research programs" in a manner which is at least partly reminiscent of theory-change in science (Lakatos 1970).

The second point supporting the existence of an abductive "consumer system" is this. Not only are abductive principles universal amongst humans, but it is hard to see how they could be other than substantially innate (Carruthers 1992). Because these principles are amongst the basic principles of learning, they cannot themselves be learned. And neither are they explicitly taught, at least in hunter-gatherer societies. While nascent trackers may acquire much of their background knowledge of animals and animal behavior by hearsay from adults and peers, very little overt teaching of tracking itself takes place. Rather, young boys will practice their observational and reasoning skills for themselves, first by following and interpreting the tracks of insects, lizards, small rodents, and birds around the vicinity of the campsite, and then in tracking and catching small animals for the pot (Liebenberg 1990). Nor are abductive principles taught to younger school-age children in our own society, in fact. Yet experimental tests suggest that children's reasoning and problem-solving is almost fully in accord with those principles, at least once the tests are conducted within an appropriate scientific-realist framework (Koslowski 1996). This is in striking contrast with many other areas of cognition, where naïve performance is at variance with our best normative principles. (For reviews see Evans & Over 1996; Stein 1996.)

In addition to a domain-general factual memory system, then, I have suggested that there may well be a domain-general faculty of abductive inference. So there are, it seems, at least two likely domain-general consumer-systems for LF representations, which either co-evolved with language, or which were specially designed by evolution at some point after language had taken on its role as the medium of inter-modular integration.

### 6.4. LF and mind reading

There are a number of reasons for thinking that the language faculty and our mind reading (or "theory of mind") faculty will be intimately connected with one another. First, there is no question but that mind reading is vitally implicated in the processing and interpretation of speech, especially in its pragmatic aspects, including such phenomena as metaphor and irony (Sperber & Wilson 1986/1995). Second, there is good reason to think that the evolution of the two faculties will have been intertwined in a kind of evolutionary "arms race" (Gómez 1998), and that language is one of the crucial inputs for normal mind reading development in young children (Harris 1996; Peterson & Siegal 1998). Finally, our mind reading faculty – specifically our capacity for higher-order or meta-representational thought – will be crucial to the operations of the sort of serial, conscious, language-using level of mentality discussed in sections 4 and 6.2 above (Perner 1998; Sperber 1996).

Almost everyone now accepts that our mind-reading capacity comes in degrees, developing in stages between nine months (or earlier) and around four years of age; and a number of related proposals have been made concerning the stages which young children pass through (Gopnik & Melzoff 1997; Perner 1991; Wellman 1990). Those who want to claim that language is implicated in mind-reading capacities would, I think, restrict their claims to the sort of meta-representational mind reading of which children become capable at about four (Segal 1998; de Villiers 2000). That is, what they find attractive is the suggestion that two- and three-year-old psychology (of the sort which we may well share with chimpanzees, perhaps) is independent of language, whereas full-blown theory of mind (four-year-old psychology or ToM) partly depends upon it. But we should distinguish between two different claims here, one of which is very likely true, yet the other of which is probably false.

The first (weaker) suggestion is just that full-blown ToM needs to access the resources of the language faculty in order to describe the contents of (some of) the thoughts being attributed to self or other. (I say this is weaker because the concept of thought in general, as a representational state of the agent which can represent correctly or incorrectly, can still be held to be independent of language – see below.) And if you think about it, something like this must be true if any version of the present proposal about the role of language in linking together different modules is correct. For if geometry and color (say) can't be combined in a single thought without language, then one could hardly expect the mind-reading faculty to be able to *attribute* a thought to another (or to oneself) which conjoins geometry with color without deploying language! This would be to give that faculty almost-magical super-properties possessed by no other module. No, if *entertaining* the thought that the object is to the left of the blue wall requires tokening the LF representation, "The object is to the left of the blue wall," then *ascribing to someone the belief that* the object is to the left of the blue wall would similarly require the use of that LF sentence.

There may be a more general point here, which gives the element of truth in so-called "simulationist" theories of our mind-reading abilities (Goldman 1989; 1993; Gordon 1986; 1995; Heal 1986; 1995). For in order for you to know what someone is likely to infer from a given thought, you will have to deploy your general – non-ToM – inferential resources, including any that are modular in nature. Otherwise we will have to think of the mind-reading system as somehow encompassing all others, or as containing a meta-theory which describes the operations of all others. This is in fact the reason why many – including Nichols and Stich, 1998, and forthcoming; Botterill and Carruthers 1999; and others – are now defending a sort of mixed theory/simulation view of our mind-reading abilities.

The second, stronger, hypothesis would be that the very concept of thought is dependent upon language, and requires an LF vehicle. According to this view, you can only have the concept of belief, say, as a representational and potentially-false state of an agent, if you are a language-user – specifically, if you have mastered some version of the that-clause construction made available in all natural languages (Segal 1998). It is this hypothesis which I would be inclined to deny. For there is evidence from aphasic adults, at least, that people who have lost their capacity for mentalistic vocabulary can nevertheless pass false-belief tasks of various sorts (Varley 1998). So I think that the full, four-year-old, ToM system is a language-independent theory which comes on line at a certain stage in normal development (albeit with that development being especially accelerated by the demands of interpreting linguistic input – Harris 1996; Peterson & Siegal 1998), which nevertheless has to access the resources of other systems (including the language faculty) in order to go about its work of deducing what to expect of someone who has a given belief, and so on. But of course the question is an empirical one, and the stronger hypothesis may well turn out to be right.

One way of pointing up the difference, here, is that on the weaker hypothesis you can do false-belief without language (at least in intra-modular contexts), although there will be many thoughts which you will be incapable of attributing (namely, those which are dependent upon language); whereas on the stronger hypothesis the capacity to solve false-belief tasks will be language-dependent across all contexts. Another empirical difference between the two accounts would appear to be that on the weaker hypothesis false-belief tasks which deal with contents drawn from a single module should be easier than those dealing with cross-modular contents – for the latter but not the former will need to operate on an LF representation. But on the stronger hypothesis there should be less or no difference, since *all* higher-order thoughts will deploy LF representations. These predictions cry out for experimental investigation.

## 7. Future evidence

What further sorts of evidence would either confirm, or disconfirm, the hypothesis that natural language is the medium of cross-modular thinking?

One obvious way forward would be to undertake many more dual-task studies of the sort conducted by Hermer-Vazquez et al. (1999). Subjects might be asked to solve problems which require information to be conjoined from two or more central-modular domains – for example: geometrical and folk-biological, or geometrical and folk-physical, or folk-biological and folk-physical. They might be asked to solve these tasks in two conditions – in one of which they are asked to shadow speech (hence tying up the resources of their language faculty), and in the other of which they are asked to shadow a rhythm (tying up their working memory to an equal or greater degree). If they fail on the task when shadowing speech but not when shadowing rhythm then this would be further evidence in support of the thesis that it is language which is the medium for integrating knowledge from different conceptual modules.

If it were to turn out that subjects perform equally well in the two conditions (either succeeding in both or failing in both), then would this be evidence *against* our thesis? Perhaps. But some caution would need to be shown before drawing any such conclusion. For one difficulty standing in the way of developing such studies is that of ensuring that the tasks genuinely do involve the *conjoining* of information across domains, and that they cannot be solved by accessing that information sequentially. So if subjects succeed under each of the conditions (whether shadowing speech or shadowing rhythm) in a task requiring them to tap into both folk-physical and folk-biological information, say, this might be because they first use information from one domain to make progress in the task and then use information from the other to complete it. (One of the distinctive features of the problems chosen for study by Hermer-Vazquez et al., is that it was known independently that in conditions of spatial disorientation, geometric information is relied upon exclusively in the absence of language.) Some care will therefore need to be taken in designing the relevant experiments.

In addition, of course, it is still to some degree controversial whether (and if so which) central modules exist. So a negative result in a dual-task experiment might be because one or more of the supposed conceptual modules chosen for the study doesn't really exist, not because language isn't the medium of inter-modular communication. But as often has to happen in science, we can regard such dual-task studies as jointly testing *both* the thesis of conceptual modularity *and* the claim that language is the medium of inter-modular integration. A positive result will count in favor of both claims; a negative result will count against one or other of them (but not both).

In principle, dual-task studies might fruitfully be devised wherever researchers have independent reason to believe in the existence of a conceptual module. For example, those who believe that there is a special system for recognizing and figuring out the degree of relatedness of kin, or those who believe that there is a special system for processing social contracts and detecting cheaters and free-riders, might construct a dual-task study to test whether the conjoining of information from these domains with others requires language. But I want to stress that such studies should *not* be conducted in the domain of mind-reading (despite the fact that many believe in its central-modular status), because of the points made in section 6.4 above. Since there is reason to think that mind reading routinely co-opts the resources of the language faculty in any case, failure in a speech-shadowing task but not in a rhythm-shadowing task would not necessarily count in favor of the thesis that language is the medium of cross-modular thinking.[13]

The other obvious place to look for evidence for or against our thesis, is in connection with either global or agrammatic aphasia. If subjects who are known to lack any capacity for formulating natural language sentences fail at tasks requiring them to conjoin information from different conceptual-modular domains, but can pass equivalently demanding tasks within a given domain, then this would be strong evidence in support of our claim. (This is a big "if," of course, given the difficulties of discriminating between patients who have lost all grammatical competence and those who have problems only with linguistic input and output.) And here, as before, if such subjects should turn out to pass both types of task, then this will count either against the thesis that language is the medium of inter-modular integration; or the existence of one or other of the supposed modules in question (but not both).

## 8. Conclusion

This paper has reviewed a wide range of claims concerning the cognitive functions of language. At one extreme is the purely communicative (or input-output) conception of language, and at the other extreme is the claim that language is required for all propositional thought as a matter of conceptual necessity, with a variety of positions in between these two poles. Section 2 discussed some versions of the cognitive conception of language which are too weak to be of any deep interest; and section 3 considered some claims which are too strong to be acceptable. Section 4 expressed sympathy for a variety of "dual process" models of cognition, especially the claim that language is the vehicle of conscious-conceptual thinking. But it pointed out that to be plausible (given the truth of central-process modularism) such views must depend on the prior and more fundamental claim that language is the medium of cross-modular thought. This then became the focus of our enquiries in sections 5 through 7, where evidence was adduced in its support, its implications discussed, and a call for further experimental testing was posted.

In closing, however, let me provide a reminder of the *character* of the exercise we have undertaken. Almost every paragraph in this paper has contained claims which are still controversial to some degree, and yet there hasn't been the space to pursue those controversies or to defend my assumptions. This has been inevitable, given the array of theories we have considered, and the range of considerations and types of evidence which are relevant to their truth, drawn from a variety of academic disciplines. But then the task has only been to survey those theories, and to show that some of them are well enough motivated to warrant further investigation – not to nail down and conclusively establish a precisely formulated thesis. And in that task, I hope, the target article has succeeded.

ACKNOWLEDGMENTS
My thanks to Colin Allen, José Bermudez, George Botterill, Daniel Dennett, Edouard Machery, David Papineau, Josef Perner, Michael Siegal, Liz Spelke, and an anonymous referee for *BBS* for their comments on an earlier draft.

NOTES

**1.** Philosophers and logicians should note that Chomsky's LF is very different from what *they* are apt to mean by "logical form." In particular, sentences of LF don't consist just in logical constants and quantifiers, variables, and dummy names. Rather, they are constructed from lexical items drawn from the natural language in question. They are also syntactically structured, but regimented in such a way that all scope-ambiguities and the like are resolved, and with pronouns cross-indexed to their referents and so on. Moreover, the lexical items will be semantically interpreted, linked to whatever structures in the knowledge-base secure their meanings.

Note, too, that an appeal to LF isn't strictly necessary for the purposes of the main thesis of this article. I use it more by way of illustration, and for the sake of concreteness. All that is truly essential is that there should exist a separate mental faculty for processing natural language, with both input and output functions, and that this faculty should deal in structured representations.

**2.** Admittedly, developmental psychologists have – until very recently – tended to down-play the significance of testimony (and hence of language) in child development. Following Piaget, they have mostly viewed children as individualistic learners – acquiring information for themselves, and developing and testing theories in the light of the information acquired (e.g. Gopnik & Melzoff 1997). See Harris (2002) who makes a powerful plea for the role of testimony to be taken much more seriously in accounts of child development.

**3.** This is, in fact, a weak version of the *Whorfian hypothesis,* to be discussed in its strongest form in section 3.2.

**4.** This date for the first appearance of fully-syntactic natural language seems to be quite widely adopted amongst cognitive archaeologists – see Mithen 1996 – so I, too, propose to accept it (albeit tentatively) in what follows. But it is, of course, still highly controversial. And it should be noted that at least some of the evidence for it turns on assumptions about the cognitive role of language.

**5.** Here and throughout the remainder of this article I shall use the term "module" loosely (following Smith & Tsimpli 1995, and others) especially when talking about central-process, or *conceptual,* modules. (Another option would have been to use the stylistically-barbaric term "quasi-module" throughout.) While these systems might not be modular in Fodor's classic (1983) sense – they will not have proprietary inputs, for example, and might not be fully encapsulated – they should be understood to conform to at least some of the main elements of Fodorian modularity. As I shall henceforward understand it, modules should be innately channeled (to some significant degree) and subject to characteristic patterns of breakdown; their operations might be mandatory and relatively fast; and they should process information relating to their distinctive domains according to their own specific algorithms.

**6.** Note that the computer programme *ChimpWorld* which successfully simulated chimpanzee behaviors and social structures without deploying higher-order thoughts nevertheless did employ structured propositional representations (Hughes 1993, reported in Povinelli 2000).

**7.** What is the status of arguments that take the form, "It is very hard to see how otherwise?" Do they merely reflect a lack of imagination on our part? Perhaps. But a more sympathetic gloss is that these are just standard arguments *to the best available explanation.* All theorizing, of course, in whatever discipline, has to work with those theories which can be imagined, or thought of, to explain the data. And it is often the case that there is only one theory which can be thought of to explain a given set of data.

**8.** How does this square with the data mentioned earlier, that there is nevertheless a substantial correlation between language ability and theory of mind in young children? Well, first, the Hughes and Plomin finding is that the genes for theory of mind and for verbal intelligence are not wholly independent of one another. And second, a quarter of the variance in theory of mind ability comes from the environment: and this may well be linguistically mediated in one way or another.

**9.** Note that geographical information isn't the same as geometric information; neither do the kinds of fact in question require integration with geometry. (Knowing that Paris is the capital of France doesn't need geometry.) So the finding that recall of geographical information is independent of language isn't inconsistent with the thesis that language is necessary to integrate geometric information with information of other kinds.

**10.** I follow the usual convention of using small capitals for Mentalese expressions, quotation-marks for natural language expressions, and italic to designate contents (as well as for emphasis).

**11.** In order to do this, the language faculty would need the resources to know that the short wall on the right *is* the blue wall. This might be possible if the outputs of the central modules are always partly indexical, as suggested in section 5.1 above, with referring elements tied to the contents of current perception. For we can surely suppose that the contents of *perception* are integrated – one's current perceptual state will contain a representation of the spatial layout of the room together with the colors of the walls. See Carruthers (2000, Ch. 11) for an argument that the demands of practical reasoning in relation to the perceived environment require an integrated perceptual field.

**12.** I haven't been able to find from my reading any direct evidence that trackers will also place weight upon the relative *simplicity* and internal *consistency* of the competing hypotheses. But I would be prepared to bet a very great deal that they do. For these are, arguably, epistemic values which govern much of cognition in addition to hypothesis selection and testing (Chater 1999).

**13.** Other sorts of dual-task study can be imagined which would test the thesis that language is implicated in the normal operations of the mind-reading faculty. For example, subjects might be asked to solve a pictoral version of one of the standard false-belief tasks while shadowing speech (in one condition) and while shadowing rhythm (in the other). But note that, given the different versions of the proposal discussed in section 6.4 above, care would need to be taken in choosing a mind-reading task. To test the stronger of those proposals (that language is required for the very idea of false belief), any sort of false-belief task would do; but to test the weaker proposal (that language is involved in drawing inferences from attributed beliefs) a more complex series of pictures might be required.

# 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.*

## A metamodule for conceptual integration: Language or theory of mind?

Scott Atran

*CNRS-Institut Jean Nicod, Paris, France; ISR-University of Michigan, Ann Arbor, MI 48106-1248.* **satran@umich.edu**

**Abstract:** Those who assume domain specificity or conceptual modularity face Fodor's Paradox (the problem of "combinatorial explosion"). One strategy involves postulating a *metamodule* that evolved to take as input the output of all other specialized conceptual modules, then integrates these outputs into cross-domain thoughts. It's difficult to see whether this proposed metamodular capacity stems from language or theory of mind.

The argument for language as a metamodule that integrates information across cognitive domains assumes that peculiarly human conceptual capacities of categorization, reference, and reasoning are domain-specific, or modular. These modular faculties automatically (innately and through maturation) parse the flux of human experience into manageable proportions. Otherwise, the world would seem too noisy for humans to acquire such rich and complex systems of knowledge in fairly unique and uniform ways (hence not susceptible to modeling by associationist processes, connectionist or otherwise), despite wide individual variation in exposure to diverse and fragmentary experiences (the "poverty-of-stimulus" argument).

Parsing occurs through privileged (but not exclusive or proprietary) access to certain domains of stimuli, or input (e.g., mechanical movements of rigid bodies, organic relationships among species and species parts, goal-directed interactions among self-activating agents). These mental faculties (the old Descartes-Leibniz term), or cognitive modules (the newer Chomsky-Fodor notion), then process input from their respective natural domains in highly distinctive ways to produce humankind's ordinary conceptions of the world (sometimes called "intuitive ontology"). Candidates for the limited varieties of our evolved intuitive ontology include so-called folk (or naïve) physics, biology, and psychology, as well as spatial geometry and temporally sequenced numbering.

Those who accept domain specificity or conceptual modularity face Fodor's Paradox – the problem of combinatorial explosion. Thus, the fact that none of us finds it difficult to generate and comprehend a sentence or thought in which, say, germs, gods, and grease (or the kitchen sink) coexist cannot be explained by the sort of modular processing that supposedly characterizes our intuitive ontology. One strategy to solve the problem is to introduce the possibility of a *metamodule,* whose evolved task (whether as a naturally selected adaptation or concomitant by-product) is to take as input the output of all other specialized conceptual modules, then integrate these outputs into cross-domain thoughts.

Perhaps the first to propose such a metamodule was Dan Sperber. His metamodule is part of folk psychology, that is, theory of mind (ToM). The functional domain of what he calls the *metarepresentational module* is "the set of all representations of which the organism is capable of inferring or otherwise apprehending the existence and content" (Sperber 1994, p. 60). The general idea is that once you have mental states in your intuitive ontology, including mental states of others, it's a short evolutionary step (perhaps no step at all) to forming intentions and desires to manipulate the mental states of others. Indeed, this is what human communication seems to be mostly about. Although such communication is clearly mediated by a public language (the product of a language module), "the very development of public language is not the cause, but an effect of the development of communication made possible by the metarepresentational mode."

Carruthers offers a similar solution to Fodor's Paradox; only, he attributes evolved metamodular functions to the language faculty, not ToM. Carruthers's insights about the role of language in thought are enlightening (at least to me); his suggestions for further experiments to test claims about how language integrates thoughts across domains are important and should be seriously pursued. But I see nothing here to allow a decision on whether language or ToM is ultimately responsible for cross-modular integration.

Like Sperber, Carruthers maintains that "mind reading (or 'theory of mind') . . . is vitally implicated in the processing and interpretation of speech, especially in its pragmatic aspects." Moreover, "there is good reason to think that the evolution of the two faculties will have been intertwined in a kind of evolutionary 'arms race.'" In sum, "meta-representational thought – will be crucial to the sorts of serial, conscious, language-using mentality" responsible for "non-domain-specific, cross-modular, propositional thought." Carruthers harnesses new and compelling experimental evidence from Hermer, Spelke, and colleagues to support the claim that language is crucially involved in cross-modular thinking; however, he acknowledges such studies *cannot* distinguish between language and ToM as the ultimate source of domain integration because of their being so intertwined developmentally and, presumably, evolutionarily.

It's difficult to decide for one source or the other. This is because the crucial structural aspect of the metarepresentational system, embedding of mental states (in ToM), may have coevolved with syntactic recursion (in language) (Atran 2002). Short-term memory limits iterated embedding of mental states (e.g., "Peter thinks that [Dan knows that] . . .") to five or six levels (R. Dunbar); but as with center-embedding of linguistic clauses (also memory-limited to under seven levels, N. Chomsky), computational machinery allows for indefinitely many embeddings (to any apparent limit, add: "You really believe that . . ."). By giving a person more time and external memory, more embedding is interpretable in a unique and uniform way (not predicted by associationist models, connectionist or otherwise). Without this embedding process (whatever its source), whose surface product is *logical form,* there seems to be no way to build cross-modular representations in sentences or thoughts.

In the end, preference for one source over the other may depend on whose evolutionary stories one finds congenial. But these stories (so far) are *just* speculative tales consistent with natural selection, lacking evidentiary standards for eliminating indefinitely many alternative and even contrary explanations. Although Carruthers renders some of these stories as parenthetical remarks, his reasoning suggests more than there is. When he says "this makes perfectly good-ecological-evolutionary sense" and leads to "just what one might predict," there remain indefinitely many other sensible stories compatible with the data. At best, these accounts retrodict findings. In no case do they predict surprising or significant discoveries unavailable before the "evolutionary" account was proposed (this is a frequent problem with evolutionary psychology explanation, however interesting).

Finally, Carruthers may be too kind to neo-Whorfians, who argue that *particular* languages sculpt culture-specific cognitive processes. Thus, evidence on Yukatek (as opposed to English) suggesting that subjects see similarities among objects based on material composition rather than shape or function carries no deep cognitive consequences (in reasoning, reference, pragmatic use, or categorization). Stimuli were routinely simple objects (cork, pieces of paper); when more complex objects were involved, differences between Yukatek and English disappeared. For example, our field team found Itza' Maya (closely related to Yukatek) more likely to group, say, a cedar canoe (*chem*) with an aluminum motorboat (*laanchaj*) than with a cedar tree (*k'uche'*). Frequently cited differences in spatial reasoning (e.g., more cardinal than relative positioning because of lack of prepositions in the language) may not be much greater between languages and cultures than between, say, New Yorkers (lousy at cardinal positioning) and rural Midwesterners (who talk about "the north side of the barn"). Even a near-prepositionless Australian aboriginal wouldn't likely refer to the mole on the right side of the nose only in cardinal terms (which would imply change of reference with every head movement).

ACKNOWLEDGMENTS
Thanks to Noam Chomsky and Robin Dunbar for information cited.

## The collective invention of language to access the universe of possible ideas

Roy F. Baumeister[a] and Kathleen D. Vohs[b]
[a]*Department of Psychology, Florida State University, Tallahassee, FL 32306-1270;* [b]*Department of Psychology, University of Utah, Salt Lake City, UT 84112-0251.* **rfb2@po.cwru.edu kathleen.vohs@psych.utah.edu**

**Abstract:** Thought uses meaning but not necessarily language. Meaning, in the form of a set of possible concepts and ideas, is a nonphysical reality that lay waiting for brains to become smart enough to represent these ideas. Thus, the brain evolved, whereas meaning was discovered, and language was invented – collectively – as a tool to help the brain use meaning.

Carruthers's excellent, thought-provoking article clarifies the relationship between language and thought. It highlights important distinctions among language, meaning, cognition, and consciousness, none of which is entirely reducible to the other.

Baumeister (in press) proposes that natural selection shaped the human psyche specifically for participation in culture. Culture uses language and meaning to organize social life and accumulate knowledge, thereby improving success at the basic biological tasks of survival and reproduction. In this view, all living things must deal with the physical environment to satisfy basic needs. Many species use social life and social interaction as strategies for improving how they deal with the physical environment. A few species (ours, most extensively) use culture to improve how they deal with social life and with the physical environment.

Carruthers speaks of the "evolution of language," but if one wishes to maintain the strict conceptual distinctions he promotes, it is misleading to speak of language as having evolved. The brain evolved; meaning was discovered; language was invented. Language is a tool to help process meaning. Carruthers is persuasive that some thought can occur without language, but of course the ability to process complex meanings, including combining multiple concepts into novel ideas, must remain quite limited without language.

In saying that meaning is discovered, we imply that a large set of concepts and their interrelationships lies "out there" independently of people or brains. This is in some ways a radical view that is at the opposite pole of the Sapir-Whorf (Whorf 1956) doctrine that language determines thought. Sapir and Whorf promoted the thesis that different languages produced fundamentally different ways of thinking and experiencing. But in fact, most languages express the same concepts and have similar grammars. (The order of words may vary, but the basic, underlying grammatical relationships are much the same. All languages denote things in the world and their activities, properties, and relationships.) Some writers, such as Pinker (2002), use this basic sameness to argue that grammatical structure derives from brain structure. We propose the opposite: Grammar is inherent in the basic, universal structures of meaning, and the brain evolved to enable people to represent and make profitable use of it.

The argument can be appreciated most clearly with mathematics, which is a relatively pure form of meaning. Mathematical truths, such as 3 x 8 = 24, are true everywhere. Some cultures may not have mastered multiplication, but in all cultures that have mastered it, 3 x 8 = 24. Yet we do not see this as stemming from any accident of how the structure of the brain evolved. Indeed, if it becomes possible to travel across the galaxy and meet up with intelligent life in distant solar systems – intelligence that is based on brains that evolved quite independently of events on earth – we predict that 3 times 8 will still equal 24. We also predict that their languages, assuming they have developed them, will have grammars similar to those on earth, so as to express things in the world and their activities, properties, and relationships. Again, these are inherent to the nature of meaning, not a by-product of the how brain structure accidentally happened to evolve. Logical reasoning is yet another example: *Modus ponens* will still be true on the other side of the galaxy.

Thus, there is one very large set of concepts and ideas, and all languages use it. This universe of possible ideas is an important form of environment that needs to be appreciated. It is not a physical reality, although many concepts refer to things in the physical world. Thinking inevitably uses these ideas, but as Carruthers has persuasively shown, it does not necessarily use language to do so. Still, language is by far the best tool for processing meaning, which is presumably why all known human groups have developed language. Baumeister's (in press) argument is that the human brain evolved in part to enable people to make use of this universe of possible ideas, first in terms of social communication and then more profoundly by developing culture.

As Carruthers has persuasively argued, language is optional in the sense that many cognitive functions can be executed without it. But surely he would agree that language vastly facilitates these functions. Moreover, language is most valuable (and Carruthers suggests even indispensable) for the most complex and integrative mental operations. Precise mathematical calculation, logical reasoning, grammar and syntax, narrative story construction – these are among the uses of meaning that enable human culture to flourish, and they rely most heavily on language.

We have spoken as if language is a tool that individuals can invent and use, but this ascribes too much power to the individual. Language is invariably found to reside in groups. It is understandable that Carruthers neglected this aspect, because his focus was on the cognitive rather than the social functions of language, but in our view the social aspect is profoundly and perhaps inextricably intertwined with the cognitive.

On that basis, we object gently to Carruthers's phrase describing what happens when language "has done its developmental work of loading the mind with information." Putting it that way implies that language serves a one-time function. On the contrary, language's most profound use is how it continues to connect the individual to the culture. Many special features of human psychology, such as lifelong plasticity, seem particularly suited for enabling people to live in a cultural society. Language's "developmental work of loading the mind with information" is never fully done.

Language allows human culture to surpass what other species can accomplish, because it stores information in the collective. The bubonic plague may kill a third of the population, but the culture's knowledge survives, because it is held in common. Hence language should be understood as a tool used by the collective, not just of individuals. Language allows cognition to accomplish things that no individual mind could achieve, such as multigenerational accumulation and transmission of knowledge. That was probably one of the decisive advantages that led natural selection to favor it.

## Domain-generality and the relative pronoun

José Luis Bermudez
*University of Stirling, Stirling, FK9 4LA, UK.* **jb10@stir.ac.uk**

**Abstract:** The hypothesis in the target paper is that the cognitive function of language lies in making possible the integration of different types of domain-specific information. The case for this hypothesis must consist, at least in part, of a constructive proposal as to what feature or features of natural language allows this integration to take place. This commentary suggests that the vital linguistic element is the relative pronoun and the possibility it affords of forming relative clauses.

Carruthers marshals an impressive array of evidence in support of the hypothesis that natural language has the highly important cognitive function of permitting the domain-general integration of different, domain-specific types of information. The hypothesis is put forward partly as a way of making sense of certain puzzling features of the archeological record, and partly as a way of interpreting some highly suggestive experiments carried out on the emergence of domain-specific thinking in early childhood (Hermer-Vazquez et al. 1999). As Carruthers himself admits, the power and plausibility of the hypothesis is a function, not just of the degree of correlation that can be established between language-use and domain-general thinking, but more importantly of how exactly it explains the way that language makes possible domain-general thinking. We need to know not just what language does, but how it does it. In particular, we need to know what aspects of language might plausibly be thought to make possible the transition from domain-specific to domain-general thinking.

The basic tenet of the hypothesis of domain specificity is that certain fundamental types of cognitive activity are carried out by modular systems that have evolved to deal with particular types of problems and particular types of situations. Popular candidates for domain-specific modules include the interpersonal competences involved in social interactions, the basic principles about objects and their interactions that are usually collectively labeled naïve physics, and an intuitive grasp of folk biology and natural history. These modules operate on a highly selective and domain-specific set of inputs with a fixed and limited amount of background information. There is integration of information within each module, but not across modules. An example of this failure of intermodular (as opposed to intramodular) integration can be found in the archeological record. In the Middle Paleolithic, for example, we find what seem to be highly developed tool-making skills existing side by side with a subtle and advanced knowledge of the natural environment, but it is not until the Upper Paleolithic and the emergence of language that we see these two bodies of knowledge being integrated in the form of tools specially designed for dealing with different plants and animal, together with hunting strategies that are tailored to the habits of specific animals (Mithen 1996).

Let us suppose, with Mithen and Carruthers, that language is required for the integration of domain-specific modules. In line with our opening question, we need to ask what feature of language could make this integration possible? I will suggest that the crucial linguistic phenomenon is provided by the linguistic mechanisms of quantification and the relative pronoun (the natural language equivalent of the bound variable of quantification).

The crucial feature of the relative pronoun is that it permits the formation of relative clauses. One way of thinking about relative clauses is as a way of distinguishing within a sentence between the object that is the logical subject of the sentence (what the sentence is about) from what the sentence says about that object (Quine 1974, p. 24). In English, for example, from a sentence such as "the red deer comes to the water just before nightfall" we can extract the relative clause "that comes to the water just before nightfall." This relative clause can be used to characterize other animals, or be embedded in further sentences, and so forth. Once the relative clause has thus been constructed and detached from the original sentence, the information it provides is available in a form that can be employed in a variety of different contexts.

Consider a simplified model of domain-specific cognition in which a thinker can think about two separate domains, each of which contains two objects and four properties. The properties from one domain can only be applied in thought to the objects in that domain but no object or predicate features in more than one domain. Suppose that *a* is an object in the first domain, and *G* a predicate in the second domain. How might such a creature come to be capable of the thought that *Ga*? Only by forming a conception of what it is for something to be a *G* that is no longer tied to the objects in the second domain. That is to say, only by forming a conception of what it is for an arbitrary object to be a *G*. And it is this that the introduction of the relative pronoun makes possible (Bermudez 2003). The relative clause allows complex predicates to be extracted from complete thought in a form that will allow them to be applied to objects falling under other cognitive domains, and the defining feature of domain-general cognition is that objects from one domain can be thought about in terms formerly associated only with objects from another domain, as in totemic art when an artifact is attributed the properties of an animal, or as in advanced tool construction when the design of a tool is specifically tailored to properties of the intended prey.

How does this apply to the Hermer-Vazquez experiments? As Carruthers points out, the representations required to carry out this task involve multiple embedding of predicates and phrases. We can view this as a multiple embedding of relative clauses. Consider Carruthers's sentence/thought: "The toy is in the corner with a long wall on its left and a short blue wall on its right" (sect. 5.2). Perhaps this should be parsed: "The toy is in the corner that is bounded on its left by a wall that is long and that is bounded on its right by a wall that is short and that is blue." The five relative pronouns in this sentence introduce relative clauses that qualify (a) the corner, (b) the left wall, (c) the corner, (d) the right wall, (e) the right wall. The integration of domain-specific information occurs because these relative pronouns serve to bind together into a single sentence/thought a range of relative clauses conveying information from different domains.

## Language in the modular mind? It's a no-brainer!

Derek Bickerton

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

**Abstract:** Although Carruthers's proposals avoid some of the more obvious pitfalls that face analysts of the language-cognition relationship, they are needlessly complex and vitiated by his uncritical acceptance of a highly modular variety of evolutionary psychology. He pays insufficient attention both to the neural substrate of the processes he hypothesizes and to the evolutionary developments that gave rise to both language and human cognition.

I found much of Carruthers's article quite helpful. On some issues he's surely right, and he's shown me where my own thinking has been insufficiently explicit (after all, isn't that what philosophers are for?). He's right in thinking that language is not just an input-output system for communicating ready-made thoughts, and in rejecting the "no thought without language" position. The cognitive functions of language lie somewhere between these extremes. But precisely where?

He rejects my take on their location (Bickerton 1990; 1995) – that "off-line thinking," requiring manipulation of a secondary representational system created by language, is therefore unavailable to other species – for inadequate reasons. First, he cannot see how *Homo erectus* and archaic *Homo sapiens* could have survived "in extremely harsh tundra environments" without language. Leaving aside whether any pre-Neanderthals actually did this, how does he think bears or caribou survive? Second, he finds my position inconsistent with "central-process modularism" (by which he seems to mean the brand of evolutionary psychology advocated by, e.g., Barkow et al. [1992]). That is true but irrelevant; I simply don't buy this view of the mind.

The crunch comes when he claims that "an ability to mentally rotate an image of the stone product that will result if a particular flake is struck off" (required, he believes, for the making of Achuleian hand-axes) is "surely 'off-line' thinking if anything is!" Clearly I haven't been explicit enough. What I meant by "off-line thinking" was not the capacity to manipulate mental images of preexisting objects when the raw material for those objects was physically present. Rather, it was the capacity, unique to our species, to conceive and subsequently create new objects – for example, barbed missiles (arrows, harpoons, etc.). At Olduvai, *Homo erectus* produced the same tool for a million years (Jellinek 1977). Imagine *Homo sapiens* doing that! What most sharply distinguishes us from other species is complex language and a capacity to produce an unlimited variety of both behavioral and artifactual *novelties.* These two things surely did not evolve in isolation from one another; how they are related should be the central issue of cognitive studies.

Off-line thinking isn't just imagining different versions of what's present – it's assembling and reassembling internalized representations to produce wholly new concepts, behaviors, and artifacts. Barbed missiles could not have been produced until someone conceived of "a missile that, when it struck its target, could not easily be dislodged." To generate such concepts you need, most crucially, syntax (the importance of which Carruthers grants but never really explains); note that the expression of the barbed-missile concept is biclausal, as any complex thought surely must be: if *x* then *y, x* because *y, y* whenever *x,* and so on. If, perhaps for reasons given in Bickerton (2002), protolanguage (and any extra thinking it enabled) was limited to equivalents of single clauses, the simultaneous emergence of syntax and our species provides a possible explanation for the colossal cognitive differences between ourselves and other animals.

Carruthers downplays these differences. He rejects Tomasello's (1999) compelling argument that there was insufficient time for more than one or two novel cognitive capacities to evolve in the few million years of our independent evolution. He rejects it because "only 10,000 years separate polar bears and grizzlies"! But we're not talking about minor differences such as distinguish bear species – we're talking about what Maynard Smith and Szathmary (1995) characterized as one of the major transitions of evolution, which led at least one paleoanthropologist (Tattersall 2002) to declare his belief in saltational events. This isn't normal speciation, and it involves more than time shortage – each of Carruthers's suite of cognitive capacities would then demand its own selectional history. But evolutionary psychology, despite its title and its protestations, has produced little beyond a handful of just-so stories lacking precise evolutionary whens, wheres, and whys.

The modular mind model has other shortcomings. Carruthers's article barely mentions the brain that, presumably, generates all the phenomena under discussion. Except for two paragraphs in section 6.2, little in the article would be incompatible with full-on Cartesian dualism. None of the many brain-imaging studies of the last couple of decades is referenced. This is unsurprising; little of that literature supports modularity. Indeed, it suggests that, outside the familiar sensory and motor systems, little if any cortex is uniquely dedicated to a single function. Rather, numbers of neurons may be recruited for whatever cognitive task attention mechanisms currently focus on.

This picture yields at least two alternatives to Carruthers's interpretation of experiments that show differences between solving orientation problems while shadowing speech and shadowing rhythms. Rather than the first task being hampered because language mediates between cognitive modules, the first task may recruit more neurons than the second, or the first task may involve circuits shared by both the functions concerned, whereas the second does not.

Indeed, Carruthers's concept of language as "the medium of non-domain-specific thinking" remains problematic even if one accepts the modular mind. While all parts of the brain aren't equally well connected, none involved in cognition come even close to being self-contained, either in our species or (as far as we can determine) others. So why did we need specifically language to integrate modular outputs? And exactly how did language perform the integration?

Carruthers's explanation assumes that an overwhelmingly modular brain required some "non-domain-specific central arena" to receive inputs from various modules, merge them, and organize their output. Language forms just such an arena. But why didn't some form of "mentalese" or "language of thought" perform this task? Why did the arena require a physical output through a specific modality?

The closest he comes to an explanation is to suppose that we somehow must "hear" inner speech before we can understand it and therefore must first give it some kind of phonological representation (this is the only way I can interpret the second paragraph in sect. 6.2). But who "hears" inner speech? I don't – I'd think I was having auditory hallucinations if I did – and I know of no evidence that others do. Carruthers certainly provides none. The unfortunate term *inner speech* merely renames the Joycean stream of consciousness – the largely spontaneous, unfocused cerebral activity that occupies much of our waking life. We need to remember that there is no *speech* in the brain, no *images,* no *sentences,* no *thoughts* – only chains of electrochemical impulses that, depending on their locations and/or routing, may appear to us in one of these guises.

The concept of a language module busily inputting, outputting, splitting, and cohering material from many quasi-autonomous cognitive modules creates a needlessly complex (as well as neurologically ill-supported) model of how language affects cognition. It's more parsimonious and more congruous with what we know of the neurophysiology of language (see, e.g., a good recent review such as Pulvermuller [2002]) to suppose that its semantic component, present in protolanguage, provides a layer of representations additional to those found in the brains of other species, whereas its syntactic component, unique to true language, builds from those secondary representations the complex cognitive constructs underlying our equally unique behavioral plasticity. There's at least provisional grounding for such a proposal in the actual details of human evolution (Bickerton 2002; Calvin & Bickerton 2000). Carruthers fails to ground his model in those details. But such grounding is essential for any adequate theory of language and cognition.

## You don't say: Figurative language and thought

Gregory A. Bryant and Raymond W. Gibbs

*Department of Psychology, University of California, Santa Cruz, Social Sciences II, Santa Cruz, CA, 95064.* **gbryant@ucsc.edu gibbs@ucsc.edu**

**Abstract:** Carruthers has proposed a novel and quite interesting hypothesis for the role of language in conceptual integration, but his treatment does not acknowledge work in cognitive science on metaphor and analogy that reveals how diverse knowledge structures are integrated. We claim that this body of research provides clear evidence that cross-domain conceptual connections cannot be driven by syntactic processes alone.

The relation between language and thought has been central to research examining the role of metaphor and analogy in communication and reasoning. Carruthers's hypothesis that language may function to integrate the output of conceptual modules has a certain appeal. For example, it is consistent with the fact that language does not seem inextricably tied to one modality. Traditional approaches to the evolution and function of language generally focus on speech communication, and this is clearly not the entire picture. Carruthers also makes specific claims about which aspects of language are conscious and which are not. By distinguishing the parts of the system that function to integrate the output of conceptual modules from the communicative mechanisms (i.e., language production and comprehension, and inner speech) we can begin to make proposals about the phylogenetic history of this multifaceted capacity, and design experiments that tease these elements apart.

However, Carruthers's treatment ignores cognitive science research that has explicitly explored the relations between language and thought. This work strongly suggests that many aspects of language, ranging from word meanings to conventional and novel expressions, reflect enduring conceptual mappings from diverse domains of knowledge and experience. A large body of work in cognitive linguistics and experimental psycholinguistics on metaphor and analogy reveals how structural mapping processes exploit common relations across domains (Gentner et al. 2001), and help structure abstract, often vaguer, aspects of knowledge in terms of concrete, often embodied, source domains (Fauconnier & Turner 2002; Gibbs 1994; Lakoff & Johnson 1999). For example, enduring conceptual metaphors like LIFE IS A JOURNEY and KNOWING IS SEEING map diverse knowledge domains to provide a rich set of entailments that are evident in everyday language (e.g., "We came to a turning point in our research project" and "I can't see what you mean in this article"). A significant body of empirical research suggests that people tacitly employ conceptual metaphors in many aspects of reasoning, memory, and language use (Gibbs 1994; Katz et al. 1997).

More recent linguistic and psychological studies demonstrate that certain enduring conceptual mappings arise from pervasive correlations in embodied experience such as KNOWING IS SEEING, MORE IS UP, DIFFICULTIES ARE PHYSICAL BURDENS, and SIGNIFICANT IS LARGE. These "primary metaphors" are not temporary mappings, perhaps created just for communicative purposes, but reflect people's embodied understandings of the positive correlations in their experiences of both the source and target domains (Gibbs et al., in press; Grady 1999). Related work on structure mapping theory suggests that structural consistency constraints and a systematicity principle guide the mapping of relations between domains dictating which sets of structures can be most readily coupled (Gentner & Bowles 2001). We claim that this body of research on metaphor and analogy provides clear evidence that cross-domain conceptual connections cannot be driven by syntactic processes alone, as argued by Carruthers. An important element of this work is that conceptual knowledge is not represented solely by propositional structures but also includes various nonpropositional information (e.g., neural nets, image schematic structures). These alternative representational formats reduce the need for resolving propositional representations by syntactic processes.

Carruthers's proposal does not specify any constraints on which domains can be easily integrated, nor does it say how this can actually be accomplished. How are different domains hooked into the system? How does evolution create a system that integrates what would be safe to assume are domain specific representational formats? The answers to these questions appear somewhat circular in that if syntax is the defining feature of language, and the LF representations act as the underlying basis of sentence production and module integration, then the module output must have syntactic structure to be exploited. We believe that a clearer, less circular explanation is needed for the grounding of conceptual relations in human thought and language.

One possibility is that structure mapping and syntactically based conceptual integration may coexist in some manner. Carruthers's model may best account for the combining of information from disparate domains that perhaps are not structurally aligned. In this way, Carruthers's hypothesis may complement research findings on metaphor and analogy. We applaud Carruthers's attempt to identify a particular functional role for language in theories of conceptual integration. But let's not forget that language often reflects enduring aspects of conceptual structure, showing how concepts from different domains are unreflectively, metaphorically contrasted and combined to allow people to make sense of their experience and the world around them.

## Language isn't quite *that* special

Joanna J. Bryson

*Department of Computer Science, University of Bath, Bath BA2 7AY, United Kingdom.* **jjb@cs.bath.ac.uk** **http://www.cs.bath.ac.uk/~jjb**

**Abstract:** Language isn't the only way to cross modules, nor is it the only module with access to both input and output. Minds don't generally work across modules because this leads to combinatorial explosion in search and planning. Language *is* special in being a good vector for mimetics, so it becomes associated with useful cross-module concepts we acquire culturally. Further, language is indexical, so it facilitates computationally expensive operations.

Carruthers has provided an excellent review of the various ways in which language can affect cognitive function, and a description of some of the most exciting new data currently available on the subject. In this commentary I will take a position largely sympathetic with Carruthers's viewpoint on the general architecture of mind, and also with his ambition of creating a "'suitably weakened'" version of Dennett's (1991) Joycean machine hypothesis. However, my position on the role of language is weaker still than Carruthers's. I do not believe language is the only way to "'cross'" cognitive modules. I also do not believe current evidence can distinguish my view from Carruthers's. To do so, I propose at the end of my commentary some experiments extending those of Hermer-Vazquez et al. (1999), described in Carruthers's section 5.2, which might help distinguish these theories.

Carruthers takes the view that modules of mind are necessarily fully disjoint, except where they are combined by language. In fact, modules may frequently interact, provided the individual modules are able to exploit each other's processing. To argue this point requires a clear definition of "*module.*" I believe that the most critical feature of modularity is that individual modules support and are supported by specialised representations. During both evolution (for the species) and development (for the individual), preferred synaptic organisations are learned for recog-}nising regularities that are useful for controlling actions and achieving goals. These regularities may arise from the external environment, as communicated by the senses, or from neighbouring neural systems – from the neuron's perspective there is no difference. This view of the development of modularity in the individual is similar to that of Bates (1999) and to some extent to that of Karmiloff-Smith (1992), whose work emphasises the developmental aspects of skill specialisation. On the species level, it echos Livesey's (1986) account of the evolution of brain organisation. These specialized representations provide the neural platform for the better-known characteristics of modularity, the specialized learning and processing, for example, of faces or language.

What are modules for? From a computation perspective, the answer is clear: Modularity combats combinatorics by focusing computational resources on a fertile but strictly limited substrate of potentially useful ideas (Bryson 2001; Minsky 1985). This explains why cross-modular processing is unusual – it is generally a bad idea. Evolution and/or learning have shown the agent that the fertile solutions for *this* problem (that addressed by one module) lie *here* (within the module). Searching other modules takes time and working memory, and is unlikely to be productive. Nevertheless, animal behaviour requires modules to interact. They interact not simply at "'centralised'" switching locations such as the basal ganglia (though that probably happens; see, e.g., Gurney et al. [1998]), but also directly. A common example is the visual system, where bidirectional channels of visual processing seem to flow through much of the cortex, with various regions specialised to particular parts of sensory processing and expectation setting (Rao & Ballard 1997). These regions qualify as modules by my criteria, because they have discrete representational maps dedicated to different levels of visual abstraction.

So modules communicate without language. I would also argue that, as elegant as the argument of the LF cycle is, it seems unlikely that language is the only module capable of perceiving its own output. First, it seems unlikely that language is composed of only one module, any more than vision is. But second, we know from Rizzolatti et al.'s (2000) "mirror neuron" work that at least one module exists in nonhuman primates that is used for both sensing and action.

What makes language special, then? It may be just two things. First, (returning to the Joycean Machine (Dennett 1991), language is tightly associated with cultural knowledge. Doing difficult computational work quickly is one of the big advantages to having a massive number of processors operating simultaneously, provided these processors can communicate their results. This is just what humans do. Thus if some individuals do manage to find cross-modular strategies that are useful, those strategies are likely to be collected in our culture and passed on linguistically. Things that look cross-modular might even be newly developed skills. It would be hard to know to what extent new skills leverage native abilities, and that extent may well vary between individuals. Spelke and Tsivkin's (2001) beautiful study appears to demonstrate that for large-magnitude approximate number values, most people utilise their innate magnitude estimation module. However, it does nothing to prove whether we ever truly learn to exploit the small exact-number module or, as we learn to count, we develop a new module that remains heavily dependent on our lexicon.

The second special thing about language is that it is indexical. That is, it is a compact way to refer to a concept. By compact, I mean that it doesn't take as much working memory, or carry as much qualia or its associated emotional and motivational baggage, as a full concept. This property of symbols has been demonstrated by Boysen et al. (1996) with chimpanzees. Two chimpanzees are presented with two bowls containing different numbers of pieces of fruit, and then one subject is asked to select the bowl the other chimp will receive. Chimps are selfish, so the goal of all subjects is to point at the smaller bowl, but the subjects find themselves incapable of this indirection in the face of the fruit itself – they always point at the larger bowl, which they want. However, they can do the task when the bowls contain numerals, which they are already trained to understand. Hauser (1999) speaks further on this task and its relationship to self-inhibition.

The conciseness of symbol-like words may explain the experiments by Hermer-Vazquez et al. (1999). Language allows the concepts of "*left*" or "*right*" to be retained in working memory (or even the phonological loop), while the other processing that needs to be done to combine navigational information is performed. We already know that learning new navigational strategies even in rats requires a resource that's also implicated in episodic memory recall: the hippocampus (Bannerman et al. 1995; McClelland et al. 1995).

There is no reason, with the possible exception of parsimony, to believe the account presented here rather than Carruthers's, or his instead of mine. However, this question might possibly be resolved using chimpanzees or bonobos. Kuhlmeier and Boysen (2002) have already been conducting experiments on chimpanzees' abilities to combine cues in navigation through using their innate ability to map from a small-scale model of their enclosure to the actual enclosure. Unfortunately, their chimps do not have any symmetric spaces suitable for the Hermer-Vazquez et al. (1999) experiment, but perhaps such a space could be constructed.

If LEFT and RIGHT happen to be concepts adult chimps acquire, they may be able to solve the task. Perhaps, if apes can learn symbols for "left" and "right", this experiment might be best conducted by one of the ape-language labs. If my indexical argument holds, then it may be that only apes who have learned symbols for these concepts can succesfully complete the task. If it is in fact the phonological loop that is critical, then it may be that only the chimps who able to carry or gesture the symbol will be able to navigate successfully.

ACKNOWLEDGMENTS
Thanks to Valerie Kuhlmeier for discussion and Will Lowe for proofreading.

## Is LF really a linguistic level?

Nick Chater

*Department of Psychology, Institute for Applied Cognitive Science, University of Warwick, Coventry, CV4 7AL, UK.* **nick.chater@warwick.ac.uk**

**Abstract:** Carruthers's argument depends on viewing logical form as a linguistic level. But logical form is typically viewed as underpinning general purpose inference, and hence as having no particular connection to language processing. If logical form is tied directly to language, two problems arise: a logical problem concerning language acquisition and the empirical problem that aphasics appear capable of cross-modular reasoning.

Let us assume, with Carruthers, that there is an internal language in terms of which domain-general reasoning is conducted (following Fodor [1975], call it the language of thought – LOT). Natural language must clearly interface with such an internal language: That is, language understanding must involve translating from natural language to LOT, and language production must involve the reverse.

What might LOT be like? The standard assumption is that LOT must be a logical language, that is, a language over which logical (or more precisely proof-theoretic) operations can be defined (see Fodor 1975; Fodor & Pylyshyn 1988). Accordingly, for each natural language sentence, there will a corresponding a "logical form" – in the traditional formal semantic sense. (Typically, there will be several such logical forms, because of ambiguities in natural language; and the picture is further complicated by anaphora, deixis, etc., but in ways irrelevant to the present argument.) The project of formal semantics (Dowty et al. 1981) seeks to explicate the relationship between natural sentences and their putative logical forms; and logical form has, of course, been a central notion in analytic philosophy and linguistics. The advocate of LOT typically takes a "psychologistic" view of logical form – logical forms correspond to representations of sentences in LOT. (One might query whether inference can be captured by logical reasoning alone – but this is another issue; see Oaksford & Chater [1991; 1998].)

Carruthers points out that Chomsky (1995) argues that there is a level of linguistic representation (called *logical form,* or LF). He then claims that all domain-general propositional thought is framed in terms of these LF representations. It is in this sense, he argues, that language has a cognitive function: a *linguistic* level of representation, LF, that plays an essential role in domain-general thought.

The interest of this claim seems to hinge on terminology: It seems to be crucial that logical form is viewed as a *linguistic* level of representation. Then, one may say that thought has, in a sense, a linguistic basis, and therefore that language has a cognitive function. If, however, we described LF as, say, a logical level of representation, then there would be no tendency to claim that thoughts (framed in LF) have a linguistic basis; or that language thereby has a cognitive function.

Carruthers's assumptions concerning LF are, as far as I can see, precisely the standard ones described above – natural language is translated in and out of LF; and LF is the representation over which domain-general inferences are defined. Thus, LF seems not to serve as a linguistic level of representation – but as a level of representation for general cognition; although one that must interface with natural language, as it must with perceptual input systems and motor output systems. In particular, from what Carruthers says about LF, it seems quite conceivable that, for example, LF developmentally and evolutionarily precedes the development of external natural language. As it happens, Chomsky (1995) has a specific theory of LF, which differs in a number of ways from standard views of logical form, but features specific to Chomsky's notion are not discussed in, and are therefore presumably not critical to, Carruthers's argument.

In a nutshell, advocates of LOT typically think of LOT as expressing the logical form of natural language sentences and thereby serving as a basis for inference. Carruthers does not appear to add substantive assumptions that establish any further sense in which logical form is fundamentally a linguistic style of representation. But without such assumptions, it seems inappropriate to draw conclusions about the putative cognitive function of language, which might more neutrally be viewed as concerning the putative cognitive functions of LOT.

One could imagine strengthening Carruthers's claim in the following way: Suppose that logical form representations can only be entertained if (at least in principle) the cognitive system can derive syntactic and phonological representations corresponding to the logical form representation (which, in principle, might mean something like "given appropriate cognitive resources"). Then the linguistic abilities of language users (their phonological and syntactic capacities) will strikingly constrain their powers of domain general reasoning. According to this view, logical form representations will necessarily be constrained, in evolution and development, by the development of language abilities; people with impaired linguistic abilities will necessarily have impaired general reasoning abilities; and so on. By making this strong connection with other levels of linguistic representation, the linguistic character of logical form representation is given real substance.

But this viewpoint, though substantive, seems difficult to defend. One problem is logical. As Fodor (1980; 1981) points out, if thought is bounded by language, then language learning is impossible. This is because language learning requires framing hypotheses about the meaning of linguistic forms that have not yet been mastered – and this requires being able to entertain meanings for which a current linguistic representation is not available. But this is precisely what is precluded on the view just described: that logical form representations are only available where the corresponding phonological and syntactic representations can be constructed. A second empirical problem is presented by patients with severe aphasia, who can carry out numerous complex tasks. For example, as Carruthers notes, an aphasic studied by Varley (1998; 2002) can, among many other things, pass the false belief task. From a modularist's perspective, this requires at least the integration of visual modules (viewing and understanding the experimental setting) and social reasoning modules (concerning beliefs and related notions) and hence is a paradigm example of cross-modal reasoning.

One might, though, defend the claim that *conscious* propositional thought necessarily implicates language. Indeed, it seems to me that this claim is almost certainly right. Conscious awareness seems to be limited to perceptual inputs (broadly construed to include awareness of bodily state, motor activity, pain, etc.) and their imagistic analogues (visual, auditory images). The only way of turning propositional thought into a form that allows it to be perceived (or imaged) appears to be to turn it into a linguistic form. Natural language is arguably the only means we have of externally representing propositional thought; if we can only be conscious of that which can (potentially) be externally represented (i.e., perceived and imaged), then natural language must inevitably underpin conscious thought.

## Anchors not inner codes, coordination not translation (and hold the modules please)

Andy Clark

*Cognitive Science Program, Indiana University, Bloomington, IN 47405.*
**andy@indiana.edu**

**Abstract:** Peter Carruthers correctly argues for a cognitive conception of the role of language. But such a story need not include the excess baggage of compositional inner codes, mental modules, mentalese, or translation into logical form (LF).

In this well-researched and very valuable contribution, Peter Carruthers defends a specific version of what he calls the "cognitive conception of language," in which language plays some direct and important role in at least some aspects of human thinking and reasoning. I think the cognitive conception is correct and immensely important. But the starting points of the story are open to both conceptual and empirical challenges. A different kind of starting point may lead to an even more interesting view of language and its role in human thought.

Carruthers's central claim is that natural language (and more specifically, the structured logical form internal representations of natural language sentences) is the medium for intermodular and non-domain-specific thinking. The structured LF representations of the language faculty provide the common format for integrating information from multiple conceptual modules, each of which might use its own form of mentalese. According to this view, thoughts are translated from mentalese into natural language representations, so that input from multiple modules can be represented in a single, integrated linguaform representation.

But suppose you are not a friend of large-scale mental modularity, and you do not believe in mentalese. Perhaps you do not even believe that the internal representations of sentences *themselves* have logical form or constitute a compositional inner code. Could you still accommodate the impressive results of Hermer-Vazquez et al. and preserve a strong cognitive conception of language?

I think the answer is yes. One possibility is to think of perceptually encountered or recalled words and sentences as acting less like inner data structures, replete with slots and apt for combinatoric action, and more like cheap ways of adding task-simplifying structure to the perceptual scene. Words and sentences, in this view, act as stable anchor points around which complex neural dynamics swirl and coalesce. Instead of thinking of linguistic encodings as enabling informational integration by acting as a common format for the outputs of multiple modules, we can think of the whole process as one not of translation into a single unifying representation but of attention-based coordination. Words and sentences here serve as kinds of simple, cheap, quasi-perceptual marker posts, enabling the agent to attend to specific dimensions of a scene, including specific combinations of aspects of the scene that would otherwise remain unnoticed.

We can also delete the modules. By all means, allow that human learning is sculpted by some innate biases, but do not suppose that the thinkings thus supported are architecturally isolated from each other. Into this nexus, learning words (such as "blue") and phrases (such as "to the left of") may be seen as the developmental source of new forms of selective attention. Learning the words (which now act much like cheap behavioral targets for reward and reinforcement routines) shapes the child's attentional biases in language-specific ways that then promote new forms of problem-solving, such as the use of conjoint geometric and color cues. And certainly, there is ample evidence that children show attentional biases that appear sensitive to the language they are learning (or have learned) – for example, Bowerman and Choi (2001); Lucy and Gaskins (2001). Smith (2001) explicitly suggests that learned linguistic contexts come to "serve as cues that automatically control attention" (p. 113).

Carruthers might (again) reply that this is just to fall back on a weak conception of the role of language in which word learning merely sculpts what is basically a nonlinguistic form of representation. Such a weak role is, he might add, undermined by the shadowing results reported by Hermer-Vazquez et al. (1999), in which linguistic activity in a distractor task impairs performance on (only) those tasks requiring integration of information across domains. But if the very process of selective attention to a complex conjoined cue required (in humans) the retrieval of at least some of the relevant lexical items, the shadowing result would also be predicted. Perhaps we need to first retrieve simple quasi-perceptual placeholders (such as words) when (and only when) a task requires us to target attentional resources on complex "unnatural," or otherwise elusive, elements of an encountered scene. Such a picture is radically different from one in which the logical form of the natural language sentence provides the skeleton for a whole new compositional internal representation unifying the outputs of multiple modules.

Carruthers also innocently misclassifies my own view (Clark 1998) as a kind of weak, purely diachronic account in which language-independent, content-bearing structures do all the real work, and in which linguistic tokens merely serve to freeze specific contents for subsequent reencounter and further reflection. Although I do think this is an important role, it is not an exhaustive one. The presence of linguistic vehicles, I argue (in the same paper cited by Carruthers), is what makes possible the important phenomenon of "thinking about thinking" (or what I there call "second-order cognitive dynamics"). Only creatures who are able to make their thoughts into stable, attendable, scrutinizable objects, by explicitly vehicling them in some way, can then turn the apparatus of thinking onto the act of thinking itself.

Throughout, Carruthers draws a firm line between a "weak" role for language as a necessary condition for certain kinds of thought and a "strong" role for language as literally constitutive of the thoughts in question. But there is space for intermediate options. In particular, there is space for the idea that thoughts about thoughts are possible only because linguistic vehicling makes thoughts available as targets of reflection. Carruthers may say that this is again the weak diachronic story. But is it? If I think to myself, "I am not thinking straight today," or if I think, "My opinions about language are probably all wrong," neither of these picks out a process any *more* extended in time than, say, having the thought that grass is green. These are all individual token thoughts. Token thoughts that target other thoughts may be possible (in humans) only in virtue of the presence, at that very moment, of inner speech and rehearsal.

In sum, I think we should also consider an alternative, perception-and-attention-based picture of the role of language in thought. Such a picture is every bit as speculative as Carruthers's own. I offer it simply as a rough-and-ready sketch of one way to pursue a strong cognitive conception without embracing the full apparatus of modules, mentalese, and compositional inner codes.

## A linguistic module for integrating the senses, or a house of cards?

Rick Dale and Michael Spivey

*Department of Psychology, Cornell University, Ithaca, NY 14853.*
**rad28@cornell.edu http://www.people.cornell.edu/pages/rad28**
**spivey@cornell.edu**
**http://www.psych.cornell.edu/faculty/people/Spivey_Michael.htm**

**Abstract:** Carruthers invokes a number of controversial assumptions to support his thesis. Most are questionable and unnecessary to investigate the wider relevance of language in cognition. A number of research programs (e.g., interactionist psycholinguistics and cognitive linguistics) have for years pursued a similar thesis and provide a more empirically grounded framework for investigating language's cognitive functions.

The wider role of language in cognition is an important and interesting question, and can be studied from many directions. In some ways, the claim that language should be implicated in wider cognition seems almost self-evident. One is reluctant to approach Carruthers's general thesis with skepticism, because it seems intuitive that language is a kind of cognitive intersection for social, perceptual, and motor information. Indeed, to many researchers in the cognitive sciences, this is an important theoretical primary.

Contrary to Carruthers's assertions, however, many of these researchers do not endorse the purely communicative conception of language. The reason for this is simple: Many researchers do not hold that language is a distinct input-output module. In general, there is a prominent trend to reject the assumption that cognition is largely modular. Because Carruthers assumes modularity as a core premise, he alienates many who would be interested in the target article's argument and whose research efforts are very relevant to a cognitive conception of language. Carruthers appears to suspect this, and he asks the reader to hold in abeyance any sentiments against the assumption. However, there are two additional assumptive burdens that are borne by those who are willing to grant his premises: reification of thought and Chomsky's use of logical form (LF). In what follows, we discuss research that brings into question these three assumptions. Most important, this research offers a more empirically grounded approach to investigating the cognitive functions of language.

First, as a minor assumption, Carruthers invokes LF representations as a vehicle for propositional thought. He provides little justification for recruiting this framework beyond the fact that Chomsky has proposed it. In fact, to our knowledge, there has been little or no empirical work substantiating the psychological relevance of LF. Because Carruthers's thesis is supposed to be a psychological one, one would expect that a firmly established psychological framework for thought would be employed in his argument. However, there is no such framework, and even those who lean toward a more classical Chomskyan perspective on language seem to be acknowledging that such purely discrete structural representations are not psychologically plausible (e.g., Ferreira et al. 2002; Sanford & Sturt 2002).

A second assumption, closely related to the previous one, is Carruthers's claim that thoughts are discrete, semantically evaluable, and have component structure. Whether or not we deny him LF for want of experimental validation, Carruthers's general position that mental states have these overall properties should be questioned. Research using headband-mounted eye-tracking suggests that mental states are dynamic and continuous rather than discrete, and are only occasionally semantically evaluable with component structure. Eye-tracking has provided psycholinguists with behavioral data about how rapidly multimodal information is integrated in cognitive processes. For example, Tanenhaus et al. (1995) used eye-tracking to demonstrate that during spoken language comprehension syntactic ambiguity resolution, as well as word recognition, can be immediately influenced by visual context. Conversely, Spivey et al. (2001) demonstrated that linguistic information actively influences the temporal dynamics of conjunctive visual search. Both of these processes unfold very rapidly and probabilistically, and they appear to share their continuously updated partial (or underspecified) representations with one another.

This brings us to modularity, Carruthers's most controversial assumption. Within the cognitive sciences, the notion of modularity has been refined in various ways, and has gradually become, though not by any means to consensus, a neurophysiological postulate that has lost much of its strict Fodorian content. Carruthers knows this but still ends up with a "conceptual" modularity that finesses ongoing debate in the cognitive and brain sciences (e.g., Kingsbury & Finlay 2001). In short, it seems to us inappropriate to assume a mosaic of innate domain-specific modules when the verdict is still out. This firm commitment to conceptual modularity also leads Carruthers to some questionable speculation that should engender skepticism even from those who accept it. For example, he hypothesizes the existence of an abductive reasoning module, neglecting a recent argument (from the father of modern modularity himself) that such a module is exactly what modularity cannot offer (Fodor 2000).

So how should the cognitive functions of language be studied? Carruthers fails to mention that there are active areas of research in the cognitive sciences dedicated to the cognitive, rather than purely communicative, properties of language. For example, the fast-growing approach to language, dubbed *cognitive linguistics,* takes as a central tenet that language is inherently scaffolded by the rest of perception and cognition (Tomasello 1998). Also, for almost three decades, the interactionist and connectionist approaches in psycholinguistics have aimed to establish broad integration of information in language processing (e.g., Marslen-Wilson 1975; Rumelhart 1977). Researchers in this area seek an answer to more general questions: How fluid is the interaction between language and other systems? Do they mutually influence each other in cognitive processing? In fact, the empirical data support a view that language is very much like the "cross-modular" entity that Carruthers proposes. The problem is that visual perception, audition, motor cortex, and so forth, also seem to integrate information across modalities. If these various modalities (including language) are already in the business of interacting continuously on essentially the same level, then there is no need for an exclusive linguistic switchboard function using logical form or any other format of representation.

To summarize, there are good reasons to doubt the assumptions in which Carruthers couches his version of the cognitive conception of language. The target article's ornate network of assumptions and speculation is slightly frustrating given Carruthers's hope to establish a "factual claim" and not some "modal claim arrived at by some sort of task-analysis." He seems instead to have established a quasi-factual claim arrived at by an analysis of many assumptions and scant evidence. His proposal for a more cognitive conception of language, however, is important and should be pursued. We have discussed a few lines of research engaged in this question and would welcome Carruthers to consider them as well.

## Why not LF for false belief reasoning?

Jill G. de Villiers and Peter A. de Villiers

*Psychology Department, Smith College, Clark Science Center, Northampton, MA 01063.* **jdevil@smith.edu** **pdevilli@smith.edu**

**Abstract:** We argue that natural language has the right degree of representational richness for false belief reasoning, especially the complements under verbs of communication and belief. Language may indeed be necessary synchronically for cross-modular reasoning, but certain achievements in language seem necessary at least diachronically for explicit reasoning about false beliefs.

We could not agree more with Carruthers's central claim that language, especially syntax, allows new forms of representation and reasoning. That has been our position with regard to the development of false belief reasoning, supported empirically with both hearing four-year-olds (de Villiers & Pyers 2002) and language-delayed deaf children (de Villiers & de Villiers 2000; 2002). We argue that there is a critical property of complements in natural language; namely, they permit the representation of *false* beliefs held by another mind, and falsity is a notoriously difficult property to represent in the absence of a syntactically structured symbol system (Fodor 1975; Jackendoff 1996; Olson 1994).

With this overlap one would expect Carruthers to embrace our position with enthusiasm, but he rejects it. We are used to rejection: Linguists accuse us of being Whorfian, anthropologists accuse us of ethnocentrism, theory of mind theorists think we seized command of their module.

There are two options for false belief reasoning: Either it is propositional or it is not. We agree with Carruthers that it is. But then we can see three clear alternatives. One is that all (or most) thought is underlyingly propositional, or mentalese (Fodor 1975). The second is that there is both mentalese and LF, but only LF does the work of cross-modular thinking. That would seem to be Carruthers's position. The third is that mentalese of sufficient complexity to handle propositional attitudes would have to be virtually identical to LF (de Villiers & Pyers 1997; Segal 1998). If so, why duplicate the functions and structures? Why not assume that natural language is the medium for such thinking, especially as LF rather than inner speech? We raised that among a list of other logical possibilities for the relationship between natural language and the language of thought in this domain (de Villiers & de Villiers 2000).

Does Varley's aphasic contradict this possibility? Not necessarily, because LF could (logically) be preserved but inaccessible to the phonological input and output systems for language. Carruthers uses Varley's case study to deny that language is needed synchronically for false-belief reasoning, but that is because of his commitment to two other connected notions: (1) ToM is a module and (2) LF is only needed for cross-modular thinking. He is also tempted to say that animals have mental state representations, arguing that their "long chains" of social reasoning imply propositional mentalese. This is where our behaviorist beginnings show. We haven't seen evidence from primates or younger children that would convince us to posit both propositional mentalese *and* LF, once you allow LF to be the medium for false belief reasoning. But Carruthers needs both if he only allows LF to be the medium of cross-modular thought. It's curious, because the arguments in favor of the subtlety of syntax and semantics needed to capture propositional attitudes seem to us so much more convincing than those needed to capture "left of the blue wall"!

Carruthers has to avoid the conclusion that false belief reasoning is dependent on language if he is to keep to the claim that it is a module. So he argues that the full theory of mind system, a module independent of language, comes on line at age four. But it is *accelerated* by interpreting linguistic input, which leaves us wondering what might happen in the absence of complex linguistic input. This language-independent module would then come on line at what? 5 years? 8 years? 25 years? In addition, Carruthers states that the language-independent ToM module "has to access the resources of other systems (including the language faculty) to go about its work." Why? In particular, "mind-reading ability routinely co-opts the resources of the language faculty." Is this because it is routinely cross-modal? Maddeningly, Carruthers does not specify sufficiently which false belief tasks count as which type: The only example provided is one in which the subject believes a proposition that is itself cross-modular, "that the object is to the left of the blue wall." So, it's all very well to "cry out for experimental investigation," but only if it's clear enough to test.

Perhaps what Carruthers has in mind is that a person without sufficient language, say, a three-year-old, can imagine the false belief of another, and can token it in some system of thought but not explicitly deduce consequences or predict behaviors from it. So, logically, the theory of mind system could be language-independent. That is the kind of picture that Clements and Perner (1994) posit for their toddlers who look expectantly at the place a character will emerge premised on his false belief but then fail when asked the simple question, "Where will he look?" However, they argue that the children's expectancy might not be propositional at this point but behavioral (Dienes & Perner 1999). To answer explicitly, a propositional format must be developed. Carruthers believes that the standard false belief tasks require only intra-modular thinking, hence not natural language, though maybe propositional mentalese. But in the *development* of such reasoning, he also admits that language plays a crucial role in input and output systems.

So the difference comes down to this. Our own data are just what one would expect if the acquisition of complementation under verbs of communication and belief in language made possible the representation of the relationships between people's minds and false states of affairs, representations that were inaccessible to explicit reasoning or incomplete before. It sounds like a good idea to us to propose that something like the LF of natural language is the format for such thinking, because LF has the necessary representational richness. But we would still need to explain why LF of sufficient complexity takes time to develop. For all we know, severe aphasics might have access still to LF, but primates would not. That is not to say there are many other subtle things that can be done (even in mind reading) without LF, and it is an exciting question to ask if such things really need propositional reasoning. Much experimental and philosophical ingenuity will be required (Dennett 1983)!

## Cross-domain thinking: Common representation format or generalized mapping process?

Peter F. Dominey

*Sequential Cognition and Language Group, Institut des Sciences Cognitives, Bron, France.* **dominey@isc.cnrs.fr**
**http://www.isc.cnrs.fr/dom/dommenu.htm**

**Abstract:** In Carruthers's formulation, cross-domain thinking requires translation of domain specific data into a common format, and linguistic LF thus plays the role of the common medium of exchange. Alternatively, I propose a process-oriented characterization, in which there is no common representation and cross-domain thinking is rather the process of establishing mappings across domains, as in the process of analogical reasoning.

Carruthers proposes that cross-modular thinking consists of the integration of central-process modules' outputs by the language faculty to build logical form (LF) representations, which thus combine information across domains, and that "all cross-modular thinking consists in the formation and manipulation of these LF representations (sect. 5.1, para. 7)." I will argue that cross-domain thinking can occur without intervention of the language faculty. Rather, such thinking relies on a generalized cross-domain mapping capability. Interestingly, this type of mapping capability can operate across diverse domains, including the mapping required for performing the transformation from sentences to meanings in language processing.

In Carruthers's formulation, cross-domain thinking requires translation of domain specific data into a common format, and linguistic LF thus plays the role of the common cross-domain medium of exchange. Alternatively, we can consider a process-oriented characterization, in which there is no common representation, and cross-domain thinking is rather the process of establishing mappings or transformations across domains, as in the process of analogical reasoning.

We can gain insight into this issue of cross-domain processing from its long tradition in the sensorimotor neurosciences. Consider the problem of cross-domain coordination required for visually guided reaching to an object. The retinal image is combined with information about position of the eye in the orbit, and the orientation of the head with respect to the body to determine the position of the object in space with respect to the body. This sensory domain representation is then used to command the arm reach that should be specified in the native motor system coordinates of the individual muscles. Interestingly, Kuperstein (1988) demonstrated that this cross-domain problem could be solved without invoking common representation format but rather by constructing a direct mapping from sensory to motor system coordinates.

Can an analogous mapping strategy be used for cross-domain thinking? In response to this question, I will illustrate a form of transformation processing for the mapping of grammatical structure in language to conceptual structure and then will demonstrate how this mapping capability extends to generalized cross-domain mapping, making this point with the analogical reasoning.

A central function of language is communicating "who did what to whom," or thematic role assignment. In this context, consider the two sentences in which the open class words are labeled.

a. John(1) hit(2) the ball(3).

b. It was the ball(3) that John(1) hit(2).

Both of these sentences correspond to the meaning encoded by the predicate hit (agent, object), instantiated as labeled hit(2) (John(1), ball(3)). For each sentence, the structural mapping from open class words onto event and thematic role structure in the meaning is straightforward (123–213, and 312–213 for sentences (a) and (b), respectively). The difficulty is that the particular mapping is different for different sentence types. This difficulty is resolved by the property that different sentence types have different patterns of grammatical function words (or morphemes) that can thus identify and indicate the appropriate ⟨sentence, meaning⟩ mapping for each sentence type. Based on this mapping/transformation characterization, we suggested that nonlinguistic cognitive sequencing tasks that require application of systematic transformations guided by "function" markers would engage language-related mapping processes. Indeed, in these tasks we observed (1) language-related ERP profiles in response to the function markers (Hoen & Dominey 2000), (2) correlations between linguistic and nonlinguistic transformation processing in aphasics (Dominey et al. 2003), and (3) transfer of training across these domains (Hoen et al. 2002). These data argue for the existence of a generalized transformation processing capability that can extend across domains and is thus a candidate for a cross-domain thinking mechanism.

Within this structural mapping context, Holyoak and colleagues (Gick & Holyoak 1983) have studied the process of analogical mapping in reasoning. A classic example involves the "convergence" schema. Consider: A general must attack a fortress at the center of a town. His army is too large to approach the fort by any one of the many paths that converge on the fort. He divides his army into small groups, each converging simultaneously on the fort. Now consider: A doctor must eliminate a tumor in a patient's thorax. The doctor has a radiation beam that can destroy the tumor, but at full strength, it will destroy the intervening tissue as well. Gick and Holyoak (1983) demonstrated that subjects could use the analogical mapping to solve the radiation problem. This analogical reasoning process does not appear to rely on translation into language or a propositional representation. Rather, we can consider that it is based on mapping of the target problem onto a nonpropositional spatial image schema of the analog problem. Thus, we can consider that not all cross-modular thinking is propositional. Similarly, when physics students discover that resistor-capacitor circuits behave like the physical mass-spring systems they have studied, an analogical mapping process is triggered that yields a number of new insights. These can potentially be expressed in language, but they do not originate in any language-related format. On the contrary, the structural properties required for analogical mapping would likely be lost in the LF conversion. A related form of nonpropositional cross-domain reasoning has been well explored in the mental models paradigm by Johnson-Laird (1996).

In summary, I have the impression that Carruthers has overextended the original function of LF as an interface between language and conceptual systems. It appears implicit in Carruthers's theory that all cross-domain thinking must be propositional (or must be of the type that can be realized by the language faculty). "All" is a strong word. The cross-domain analogical mapping examples above define cases where cross-domain interaction cannot occur via a propositional LF-like data structure. Rather, these cases require mapping processes that establish the cross-domain correspondences, independent of a neutral common representation.

ACKNOWLEDGMENT

This work has been supported by the HFSP Organization.

## The role of language in the dual process theory of thinking

Jonathan St. B. T. Evans[a] and David E. Over[b]

[a]*Department of Psychology, Centre for Thinking and Language, University of Plymouth, Plymouth PL4 8AA, UK;* [b]*University of Sunderland*

**jevans@plym.ac.uk**

**Abstract:** Carruthers' proposals would seem to implicate language in what is known as System 2 thinking (explicit) rather than System 1 thinking (implicit) in contemporary dual process theories of thinking and reasoning. We provide outline description of these theories and show that while Carruthers' characterization of non-verbal processes as domain-specific identifies one critical feature of System 1 thinking, he appears to overlook the fact that much cognition of this type results from domain-general learning processes. We also review cognitive psychological evidence that shows that language and the explicit representations it supports are heavily involved in supporting System 1 thinking, but falls short of supporting his claim that it is the medium in which domain-general thinking occurs.

Carruthers proposes that language is the medium for general thought, serving to integrate the outputs of various domain-specific conceptual systems. He briefly mentions links with dual process theories of thinking. The most relevant in this context are those of Reber (1993), Evans and Over (1996), and Stanovich (1999). All of these theorists argue that humans have two separate cognitive systems with distinct evolutionary histories. System 1 evolved early, is shared with other animals, allows rapid and parallel processing of information and is independent of general intelligence. System 2 evolved late, is uniquely human, is slow and sequential, requires central working memory, and is highly correlated with general intelligence. Both Reber and Evans and Over label System 1 as implicit and System 2 as explicit.

The first question is how well do these two systems map onto Carruthers's distinctions? System 1 functions are generally regarded as domain-specific, whereas System 2 allows general-purpose abstract and hypothetical reasoning. However, Carruthers apparently presupposes that domain-specific cognition is grounded in specific innate modules, which is unwarranted. Much domain-specific cognition is the result of *domain-general* learning processes. Implicitly acquired knowledge is limited to the domain in which it is learned (Berry & Dienes 1993), but this in itself is not evidence for innate specificity. There is a high degree of biological preparedness for some learning, for example, in the human visual and language systems, but general associative learning processes can account for much of our domain-specific knowledge (Almor, in press).

We translate Carruthers's hypothesis as the claim that language provides the medium of System 2 thought. This does link with some tentative suggestions made by Evans and Over (1996). We too are struck by the fact that both language and System 2 thinking are uniquely human and agree with Carruthers that language is strongly related to domain-general thought, but we will refer to further psychological support for this view.

System 1 is responsible for much linguistic understanding and production: It is automatic in function and broadly universal in efficacy across human beings. Like other System 1 processes, language delivers its products into consciousness. We comprehend the meaning of a sentence even though we have not the slightest idea of the process by which we derived it. However, the key aspect of language for the present purpose is that it supports explicit representations. Language is the medium of explication. To refer to System 1 as implicit and System 2 as explicit is to associate language with the distinction. With language, we can formulate rules and principles that apply generally. For example, if we learn statistical principles by experience, we will apply them in a domain-specific manner (Jepson et al. 1983). If we learn them through verbal instruction, we can apply them generally (Fong et al. 1986). We can also use System 2 to reason nonconstructively, inferring, for example, a disjunction "from above" with the help of general rules without knowing which disjunction is true. In the purest case of this, we can infer by logic alone that either a proposition or its negation is true, and this capacity requires language.

In psychological research, language defines the distinction between implicit and explicit knowledge. The latter can be stated; the former cannot. System 2 thinking can be made explicit in think-aloud protocols; System 1 thinking cannot. Another key aspect of System 2 thinking is that it is volitional. System 2 reasoning performance can be directly influenced by verbal instruction. For example, people use System 2 to override System 1 belief bias when strongly exhorted to reason logically (Evans & Over 1996), although the ability to do this is strongly related to general intelligence (Stanovich & West 2000).

Is language linked also to this volitional aspect of System 2 thinking? Carruthers makes several references to the notion of "inner speech" and quotes evidence that it is a prevalent aspect of our consciousness. He refers to the position of Frankish that people can frame a sentence in inner speech "as if that sentence were true" to see whether they should believe it or act on it. This position reminds us somewhat of our own theory of hypothetical thought about possibilities and of the conditional inference that follows from it in System 2 (Evans et al., in press; Over & Evans, in press). Hypothetical thought about possibilities is a purely human activity, like advanced language use, and the two activities must be deeply connected.

Language use is thus related to System 2 thinking in a number of deep ways. The extent to which it is the *medium* of such thinking, as Carruthers claims, is debatable, however. Much explicit thinking, in our view, depends on representing and manipulating mental models (Johnson-Laird 1983; Evans et al., in press). However, the extent to which such models are linguistic in nature is unclear at present. In the version of mental model theory of Johnson-Laird and his collaborators, there is not only a negation sign but various brackets and other symbols, recently even including so-called mental footnotes (Johnson-Laird & Byrne 1991; 2002). Such features cannot simply be represented by concrete images and may be linguistic in nature. However, some theorists believe that models should be regarded as concrete and have revised the theory to deal with this (Barrouillet & Lecas 1998). It has also been shown in the domain of relational reasoning (Evans et al. 1993) that at least some linguistic information that is used to build models of spatial relationships is lost after a short delay. We also know that communication between humans relies on pragmatic factors well beyond the boundaries of language (Sperber & Wilson 1995).

In conclusion, research on the psychology of thinking supports the view that language is intimately related to domain general thinking, but falls short of providing clear support for the claim that it is the medium in which such thought takes place.

## Language, consciousness, and cross-modular thought

Keith Frankish

*Department of Philosophy, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK.* **k.frankish@open.ac.uk**
**http://www.open.ac.uk/Arts/philos/frankish.htm**

**Abstract:** Carruthers suggests that natural language, in the form of inner speech, may be the vehicle of conscious propositional thought, but he argues that its fundamental cognitive role is as the medium of cross-modular thinking, both conscious and nonconscious. I argue that there is no evidence for nonconscious cross-modular thinking and that the most plausible view is that cross-modular thinking, like conscious propositional thinking, occurs only in inner speech.

In section 4 of the target article, Carruthers suggests that auditory images of natural language sentences ("inner speech") are the vehicles of conscious propositional thinking, tentatively endorsing proposals by Dennett and myself (Dennett 1991; Frankish 1998a; 1998b; forthcoming). He goes on, however, to argue that natural language has a more fundamental cognitive role as the medium of cross-modular thinking. Now, of course, the former thesis already involves a partial commitment to the latter. If conscious propositional thinking occurs in inner speech, then conscious *cross-modular* propositional thinking will do so, too. However, Carruthers claims that we also entertain *nonconscious* cross-modular thoughts (sect. 5.1), which take the form of logical form (LF) representations rather than auditory images, and which are fed directly to systems that take cross-modular inputs. This claim is, I think, a more contentious one, and I want to question whether there is any basis for it.

Is there any evidence for the existence of nonconscious cross-modular thinking? There is little or no *behavioural* evidence for it, I think. It is true that we are capable of performing some fairly demanding tasks nonconsciously – driving, for example, or walking down a busy street. But while these tasks may draw on outputs from more than one central module, it is doubtful that they require *integration* of them into cross-modular thoughts. The activities we can perform nonconsciously are typically routine ones, requiring precise behavioural control rather than creative thinking, and are not significantly more demanding than ones that other mammals can be trained to perform. Tasks requiring creative intelligence, however, quickly evoke conscious thought.

Rather more promising evidence for nonconscious cross-modular thinking is provided by what we may call *eureka* thoughts – episodes in which the solution to a problem pops into one's head some time after one has ceased to think about it consciously (Carruthers 2000, Ch. 3). Such thoughts frequently involve conjoining ideas in new ways, and it is tempting to conclude that they must be the product of nonconscious cross-modular reasoning. Of course, if Carruthers is right, such thoughts cannot be *initially framed* as a result of cross-modular reasoning, because by hypothesis they are constructed by the speech production system mechanically combining outputs from discrete central modules (sect. 6.1). But – Carruthers may say – it is likely that they have undergone nonconscious *filtering* before issuing in inner speech. Cross-modular thoughts might be routinely passed to the abductive reasoning faculty for evaluation, with only the most promising ones eventually emerging in inner speech. There is another possibility, however. This is that cross-modular thoughts are fed directly into inner speech without filtering, and that their evaluation takes place subsequently, as the agent "hears" and reacts to them. On this view, eureka thoughts are special, not because they have been preselected for significance, but because we recognize them as important and hold onto them, whereas others are forgotten. This view is, I suggest, more consistent with the introspective data than the alternative. After all, a great deal of inner speech is simply nonsense – whims, fancies, and absurd ideas, which are instantly dismissed. Again, then, there is no compelling evidence for nonconscious cross-modular thinking – rather the opposite, in fact.

Could we elicit experimental data that would bear on the issue? What would be needed is something like a version of the Hermer and Spelke reorientation task in which subjects were distracted from thinking consciously about what they are doing. It is hard to see how this could be arranged, however. It might be suggested that we could seek the assistance of blindsighted patients – presenting them with geometric and colour information in their blind field and seeing if they could integrate it. But again is hard to see how we could test for integration of the information without stimulating the subjects to conscious thought. (Remember that blindsighted patients do not react to blind-field stimuli unless overtly prompted to do so.)

This is not conclusive, of course, and it may be that evidence for nonconscious cross-modular thinking will emerge. Even if it does, however, this would not in itself show that cross-modular thoughts can be tokened as LF representations as well as auditory images. For it may be that auditory images can themselves be nonconscious. It is plausible to think that episodes of inner speech can be *unattended,* and on some theories of consciousness this will be sufficient for them to be nonconscious. Nonconscious inner speech might nonetheless be cognitively effective – being processed by the comprehension system and made available to conceptual modules and domain-general systems.

I have argued that there is no evidence for the existence of cross-modular thinking in any form other than inner speech. I now want to outline a positive reason for thinking that it always takes that form. It is widely accepted that there is a feedback loop within the language faculty, which takes phonological representations from the speech production system and passes them to the speech comprehension system, bypassing the articulatory and auditory systems (Dennett 1991; Levelt 1989). It is this loop that supports inner speech. But if nonconscious cross-modular thinking occurs, then – assuming Carruthers is right about its language dependency – there must be additional feedback loops, which take LF representations as input and feed them back to domain-general consumer systems, such as the abductive reasoning faculty.[1] But why should such loops have developed in addition to the phonological one? After all, contents entering the phonological loop would also reach domain-general consumer systems via the speech comprehension system – and would *in addition* be available to other processes operating specifically on conscious thoughts (see sects. 4 and 6.3). Given this, what selectional pressure would there have been to develop additional loops channelling bare LF representations? They might have been marginally faster, but that is all. (Note that it is unlikely that such loops could have developed *before* the phonological one; both would have had to develop at much the same time, subsequent to the emergence of language.) The more economical hypothesis is surely that there is just one feedback loop from the language system – the one which carries auditory images in inner speech – and that it is the channel for both cross-modular thinking and conscious propositional thinking.

NOTE

**1.** In fact, it is doubtful that LF representations could be channelled *directly* to an abductive reasoning faculty in the way Carruthers suggests. It seems likely that such a faculty would operate on mental models rather than propositional representations.

## Inner speech and the meeting of the minds

William Frawley

*Departments of Anthropology and Psychology, George Washington University, Washington, DC 20052.* **frawley@gwu.edu**

**Abstract:** Four extensions of Carruthers's arguments are given. (1) Specifics of the Vygotskyan tradition can enhance his claims. (2) Linguistic relativity might be seen as variation in how logical form (LF) and phonetic form (PF)serve working memory. (3) Language aids intermodular thinking because it makes representations maximally visible. (4) Language for intermodular thinking is not hardwired but opportunistic.

Carruthers gives substance to the claim that if mind is a *computational form,* then it is in the *form of language and mind* where they come together. I offer four extensions of his points, (1) elaborating on Vygotsky, (2) recasting linguistic relativity, (3) examining representational explicitness, and (4) suggesting opportunism in language-thought relations.

Vygotsky can be generally understood as claiming that *language surfaces where minds meet:* certain forms of language, internalized as *speech for thought,* drive metathinking (Frawley 1999). Vygotsky was concerned with minds meeting in the external world and social language internalized into personal metathinking via inner and private speech. But an individual's internal subminds might also be coordinated by *speech for thought.* Language is the vehicle of computational control as minds implement executive tasks (Frawley 1997). This view evokes Frankish's arguments that language serves the supermind by aiding the acceptance or rejection of propositions – a classic mental-control function. A Vygotskyan take on Carruthers's arguments, then, is that inner speech – imagined speech as PF recycles through the architecture – is the form of metamind for on-line, cross-domain mental computation.

Several benefits accrue from using this view of Vygotsky. One is that language-thought relations emerge developmentally as speaking and thinking intertwine. Speech, as Vygotsky said, "goes underground." When Carruthers observes developmental patterns in the use of spatial vocabulary to solve geometric tasks, he invokes a long tradition of experiment and analysis in Vygotskyan theory (Vygotsky [1986] is an early statement) on the emergence of language in on-line thinking.

A second benefit concerns the form of speech for thought. For Vygotsky, inner speech is predication, and so to claim that language drives intermodular thinking, insofar as it is a point about LF, is to claim that the way a language creates predicates at the interpretive level affects informational integration. When inner speech surfaces in truncated form as overt speech for thought, *it is entirely predicative.* Personal speech for problem solving is a rich data source for how the interpretive mechanisms of language drive informational integration (Frawley 1997).

PF's critical role in the execution of speech for metathought substantiates Vygotsky's point that speech, not language, aids thinking (*Myshlenie i rech,* meaning "thinking and speech," is the original Russian). Speech for thought is really something like speech, which raises the importance of examining how speech is implicated in working memory. If acoustic and articulatory codes are activated in working memory – and if inner speech as PF is the vehicle of intermodular processing – then there ought to be an asymmetry in language/nonlanguage disruptions in working memory tasks: Speech intrusion ought to interfere in nonlanguage tasks, but not vice versa. There is some evidence to this effect. Baddeley (1999, pp. 50–51) argues that recall of visually presented digits can be disrupted by any linguistic input, even input that is not the native language of the subject, but not by nonlinguistic noise. When suppressing speech, individuals have difficulty transferring visually presented material to short-term phonological store, thus suggesting that speech is involved in crossing modes. Further careful experiments on the role of speech in working memory across modes would be quite revealing.

A second extension of Carruthers's points raises a new view of linguistic relativity. LF and PF have known patterns of variation. Would such variation imply that differences across languages in the way syntax delivers representations for interpretation, which are then used as vehicles informational integration, are thus measures of linguistic relativity? Do different languages provide working memory alternate formats for silent predication (LF) and informational ordering (PF)? Baddeley (1999, p. 53) argues that the articulatory loop in working memory is a good checking mechanism for order of information, and measures of working memory appear to be a better gauge of individual variation and performance than IQ and general problem-solving instruments. Does linguistic relativity lie in the ways languages offer their users alternate means for silent and self-directed linguistic explicitness in on-line thinking?

Explicitness leads to the third extension. If speech serves mental computation across modules, what does cross-domain mental computation require of speech? Any answer must explain how representations are made visible to one another (Frawley 2002a; 2002b; Frawley & Smith 2001). At interfaces, representations must be maximally overt, so that one domain may check the truth of the output of another. Speech is a good vehicle of intermodular computation because it accommodates degrees of representational explicitness. (Indeed, this demand for intermodular visibility is probably why visual imagery is not a good vehicle of cross-domain thought.) If Carruthers is right, intermodular knowledge disorders ought to evince ancillary, self-regulatory speech disruptions. Indeed, this seems to be the case for alexithymia, arguably a cross-domain emotional deficit whose signal symptom is a deficit in affective talk, even though the syndrome is not a language disorder at all (Frawley & Smith 2001).

Vygotskyan theory and intermodular mental computation come together in the fourth extension of Carruthers's points. Is inner speech hardwired in the modular breach? Speech for cross-domain thinking may be opportunistic rather than hardwired. Some connections across domains might become automatic and implicit with learning – as in the way an experienced navigator might develop packaged cross-domain talk to presolve the integration of geometric and nongeometric information. But novices at such tasks will have to deploy speech for thought. Language for thinking thus should be opportunistically deployable. This would explain why, in the experimental results that Carruthers quotes, speech was disruptable in environments where subjects had a wider spatial array than a narrower one. The former, more difficult, environment demands speech deployed for metathought. In less informationally rich environments – smaller rooms – less demand does not evoke opportunistic speech for thought. This point, however, is also Vygotsky's, but of some 70 years ago – that is, inner speech surfaces when tasks are difficult.

Does language help thinking? The Vygotskyan answer seems to be: only when it can.

## No conceptual thought without language

Christopher Gauker

*Department of Philosophy, University of Cincinnati, Cincinnati, OH 45221-0374.* **christopher.gauker@uc.edu**
**http://asweb.artsci.uc.edu/philosophy/gauker**

**Abstract:** Carruthers labels as "too strong" the thesis that language is necessary for all conceptual thought. Languageless creatures certainly do think, but when we get clear about what is meant by "conceptual thought," it appears doubtful that conceptual thought is possible without language.

Carruthers holds that there is *some* kind of thought that language alone makes possible, but he denies that language is necessary for specifically *conceptual* thought. He thinks that nonlinguistic creatures plainly do engage in conceptual thought, which shows that language is not necessary for it. I reply that while nonlinguistic creatures solve problems, and much of their problem-solving deserves to be called thought, it is doubtful whether anything they do deserves to be called "conceptual thought."

Here is how I would define conceptual thought. The most basic kind of conceptual thought is a thought that represents some particular as belonging to some general kind or represents several particulars as standing in some relation. For example, my thought that *that* (I am pointing at a table) is a table, is a basic conceptual thought, because it represents *that* thing as belonging to the kind *table.* We can define the rest of the class of conceptual thoughts recursively as thoughts that are either (1) basic conceptual thoughts or (2) thoughts that are *inferentially related* to other thoughts that already qualify as conceptual thoughts. For example, if from basic conceptual thoughts to the effect that *this* is a vicious dog and *that* is a vicious dog, someone infers that all dogs are vicious, then the latter is a conceptual thought; and if from the thought that all dogs are vicious, someone infers that people should not keep dogs as pets, then that is a conceptual thought as well.

We know that people who speak a language can engage in conceptual thinking in this sense because they can overtly *say* of some particular that it is of some kind, or can overtly *say* of some particulars that they stand in some relation to one another, and they can draw inferences from sayings of these kinds. For example, one person may say to another, "That's a vicious dog," and from that someone may infer, "All dogs are vicious," from which someone might infer, "People should not keep dogs as pets." What kind of evidence could we have for thinking that a creature without a language was thinking of a particular individual and representing it as belonging to some general kind?

One kind of evidence sometimes cited is the fact that nonlinguistic creatures *categorize* (e.g., Herrnstein 1992; but not Carruthers). That is, they systematically distinguish between the members of contrasting categories. For example, a pigeon might learn to peck when shown a picture of a tree (whether or not it is one of those pictures on which it was trained) and to refrain from pecking when shown a picture of something other than a tree. But this is not a persuasive basis for attributing conceptual thought. One alternative explanation of categorization might treat it as a product of imagistic thinking. A pigeon may perceive image $x$ as being more like image $y$ than like image $z$ without having to think of $x$ and $y$ as belonging to some kind, such as *tree,* and even without having to conceive of some particular respect in which $x$ is more like $y$ than like $z$.

What Carruthers actually says is that nonlinguistic creatures engage in "sophisticated" thinking, as if we could identify all "sophisticated" thinking with conceptual thought. I think we should flatly deny that some of these kinds of sophisticated thinking even do take place. Carruthers mentions "mind-reading." No doubt, people with language, even children in the early stages of language learning, attribute mental states to agents, and to that extent they engage in mind reading. But because they have language, that is not evidence for the independence of conceptual thought from language. As for languageless creatures, there does not seem to be any indisputable evidence that they are aware of other agents' mental states at all (see Heyes 1998). If people or animals use a theory of mind to predict behavior, then it should be possible to state, at least by way of hypothesis, the content of the theory they might be using. But beyond vaguely asserting that beliefs and desires somehow lead to action, no one has ever done that.

Some of the other sorts of sophisticated thinking that Carruthers alludes to surely do take place, but it is not obvious that they depend on conceptual thought. For example, I see no reason to assume that thinking that is "genuinely strategic in nature, apparently involving plans that are executed over the course of days or months" requires conceptual thought. Instead, it might involve elaborate mental movie making. The animal might imagine various courses of events and then recognize the similarity between what is actually happening and the initial segments of some of these mental movies. Carruthers thinks that Horgan and Tienson's "tracking argument" shows that nonlinguistic creatures are capable of "compositionally-structured thoughts." But sheer compositionality is no proof that thought is conceptual; imagery may exhibit a kind of compositionality as well. In Horgan and Tienson's example of the basketball player, there are features of the agent's thought that cannot be accounted for imagistically, such as the "coach's latest instructions" (Horgan & Tienson 1996, p. 85), but those are the ones that depend on spoken language.

Tradition and the contemporary majority hold that language serves communication by allowing speakers to reveal to hearers the conceptual contents of underlying thoughts. It is not clear to me that Carruthers thinks there is anything fundamentally wrong with that. He just thinks that the various aspects of the thought expressed do not come together in one place, so to speak, until the sentence of spoken language is formed. But if, as I have urged, language is the very medium of conceptual thought, so that there is no conceptual thought without language, then this traditional picture of interpersonal communication must be abandoned. The alternative is to think of language as a tool for interpersonal cooperation by which people can constrain one another's actions and steer a common course (Gauker 2003). On this conception, language can also play a role in intrapersonal thought, but that role is best conceived as the internalization of the methods of interpersonal cooperation.

## Language's role in enabling abstract, logical thought

James A. Hampton
*Psychology Department, City University (London), London EC1V OHB, UK.*
**hampton@city.ac.uk http://www.staff.city.ac.uk/hampton**

**Abstract:** Carruthers's thesis is undermined on the one hand by examples of integration of output from domain-specific modules that are independent of language, and on the other hand by examples of linguistically represented thoughts that are unable to integrate different domain-specific knowledge into a coherent whole. I propose a more traditional role for language in thought as providing the basis for the cultural development and transmission of domain-general abstract knowledge and reasoning skills.

What is the role of language in thought? Carruthers claims that thought originally developed in prelinguistic hominids and other mammals within domain specific modules such as a folk biology, folk physics, or folk psychology, and that the evolution of language was necessary for integrating the outputs of these modules and so enabling humans to be more intelligent and adaptive. I take issue with the thesis on four grounds.

First, it is unclear why language should be considered *necessary* for integrating the outputs from different domain-specific modules. The evidence cited by Carruthers is very limited for such a general claim. There are clearly unfamiliar situations in which being able to use verbal coding and inner speech will make information of different kinds easier to integrate, recall, and use, but that does not provide evidence that we cannot easily integrate information from different domains *without* language in many familiar contexts. When attempting to tackle an opponent in rugby or football, or when selecting a shot to play in tennis, the athlete needs to be able to integrate psychological reasoning about the likely action about to be taken by the other player with physical information about the speed and direction in which both players and/or the ball are travelling. It is an empirical question whether inner speech might play some role here, but it seems unlikely to be necessary for integration, given that reaction times are much too quick to allow verbally mediated conscious deliberation.

Second, it is important to note (although Carruthers does not claim to the contrary) that language may often be *insufficient* to ensure integration across domains. Although we have a well-established vocabulary with which to talk about psychology and physics, the two domains stubbornly resist attempts to integrate their outputs in any known form of language, even for experts. So a psychiatrist may feel warranted to tell her patient that there is "nothing physically wrong with you," and then proceed to successfully treat the "purely psychological" symptoms with a drug known to affect physical systems in the brain. Cartesian dualism is probably the most widely held commonsense view of the mind-body relation, and the philosophical problems of integrating the output from the two forms of discourse are notorious.

The failure of language to achieve integration may also be found between folk biology and folk physics, with commonsense beliefs in animism holding for many centuries and still demonstrable in children before exposure to science education. There are also many domain effects in reasoning. Tasks of logical reasoning such as Wason's (in)famous four-card problem are very difficult in the abstract, but can be made trivially easy if placed in a domain where the structure of the domain lends meaning to the problem. If language is successful at facilitating domain-general thinking, then an account is needed of why the task (which is presented entirely in linguistic terms) should become so much harder when abstracted away from a specific domain.

Third, the thesis ignores what may be a much more crucial role for language to play in thought, namely, *abstraction* (see, e.g., the early research of Bruner et al. 1966). More recently Barsalou (1999) has developed the thesis that (broadly speaking) concepts are mental simulations. Tokening a thought with a particular conceptual content leads to (and may even just *be*) activation of sensory and motor schemas associated with perception and action related to that concept. If his thesis is true, then thinking will operate most successfully within a modality-specific and situated form of mental model. I would argue that a key function of language is that it provides an escape from this primary level of thought. It provides the bridge between the "messy" prototype representation of a concept built around experience and action in the world and the "clean" representation of a concept as an encapsulated atom (see, e.g., Fodor 1998). Fodor's arguments for conceptual atomism involve the compositionality and systematicity of thought. Yet these arguments themselves are simply derived from earlier considerations of the compositionality and systematicity of language. Language allows us to abstract out a notion of a concept from the individual experiences on which it is based. It creates a new mental "entity" as a reified object of thought. As Dwight Bolinger put it, "The act of naming, with all we have seen it to imply in the way of solidifying and objectifying experience, becomes one of our most powerful suasive tools, enabling us to create entities practically out of nothing" (Bolinger 1975, p. 251).

Finally, I argue that a second crucial cognitive role for language is as a necessary precursor for the cultural development of socially shared *logical* thinking and hence, for rationality, narrowly defined. Natural language is, of course, a very imperfect tool for expressing truths of any kind because of the vagueness to be found in almost all of its referring expressions and the ambiguity in much of its syntax. However, once the reference of terms is established (e.g., by indexicals or stipulation), and care is taken in choice of syntactic form, the development of logic (and hence mathematics and science) can proceed, using forms of propositional argument based on the sentences of natural language. The formation of a community of speakers sharing the same language has therefore been essential for the progression of human thought through the formulation and the resolution of conflicting views.

In a sense, of course, abstract logical thinking performs the function of taking input from domain-specific modules and integrating them, just as Carruthers argues. But it is not the availability of language per se that lies at the basis of this integration. There are many examples of domain-specific knowledge that signally fail to be integrated in spite of fully developed language structures for talking about them. It is rather the fact that with language – and the cultural development and transmission of knowledge that it supports – comes the hard-won capacity for abstracting the form of an argument away from its content, developing general reasoning skills and finding the means for testing and challenging the reasoning of others.

## Relativistic implications of a natural-language-based format for thought

Steve Henser
*Independent Researcher, 174 Pennant Road, Llanelli SA14 8HN, United Kingdom.* **steve.henser@btinternet.com**

**Abstract:** I will argue (*contra* Carruthers) that accepting natural language as the format of many of our thoughts should entail accepting a version of Whorfian relativism and that, rather than something to be avoided, evidence from bilingual cognition suggests that incorporating this idea into future research would yield further insights into the cognitive functions of natural language.

Peter Carruthers's work on the cognitive functions of language is an excellent example of the healthy symbiosis of philosophy and psychology. Carruthers's work in this area has matured in response to empirical evidence from scientific psychology (cf. Carruthers 1996, where no clear theory of central cognitive modules is presented, with Carruthers 1998 and subsequent work) and, at the same time, Carruthers has injected a philosopher's objectivity into the debate and has pointed out important new areas for research.

I am particularly excited by Carruthers's identifying high-level representations of our specific natural/public languages (English, Spanish, Japanese, ASL, etc.) as being constitutive of many of our thoughts and claiming that our inner use of natural language qualitatively enhances our ongoing cognitive abilities. I feel that these claims capture an important insight, one that has the potential to motivate diverse research approaches towards a unified understanding of the language/thought relationship.

I will focus here on one aspect of Carruthers's thesis that I feel needs reworking, namely, the relativistic implications of accepting natural language (NL) as the format of many of our thoughts. In Carruthers's model, the mind contains a non-domain-specific central arena ("a very large and significant central arena" Carruthers 2003, in press), which integrates data from central cognitive modules and which, crucially, operates by manipulating the higher-level representations of specific natural languages.[1] In additional to this integrative role, NL resources are said to be co-opted by the theory of mind module and by a putative virtual faculty of theoretical reason (cf. Carruthers, in press b). Thus, a specific NL (as opposed to a universal mentalese) is seen to be the format of or constitutive of many of our thoughts, to be directly involved in key higher thought processes, and to be responsible for our distinctly human cognitive abilities. The influence of NL on thought is thus seen to be not only developmental or diachronic but also of a continuous or synchronic nature. In view of the crucial place in cognition accorded it in his model, it is strange that Carruthers feels obliged to reject the Whorfian, relativistic view of specific natural languages as exerting an influence over the way their speakers think.

From the brief dismissive mention of Whorf in the target article, it seems that Carruthers has espoused a strong deterministic interpretation of Whorf's views, but this does not do justice to their scope or complexity (cf. Lakoff 1987, pp. 328–30). Although Whorf wrote of a "linguistically determined thought world" (Whorf 1956, p. 154), he was not suggesting that humans are totally the conceptual prisoners of their first languages of acquisition. Taking an overall perspective of Whorf's work, it seems more appropriate to think of him as advocating NL as exercising both a formative influence on speakers' concepts and an ongoing influence in habitual thought. This position is surely consistent with Carruthers's thesis, at least insofar as it applies to the very extensive part of the mind that is said to operate by deploying NL resources; it is also consistent with a considerable body of work in cognitive psychology and neuropsychology that supports the existence of language-specific effects in cognition (e.g., Emmorey 1998; Hunt & Agnoli 1991; Zhang & Schmitt 1998).

Carruthers's theory, while groundbreaking, may not go far enough in pointing up the possible synchronic qualitative cognitive enhancements available to an organism by virtue of its possession of NL. Not only the combinatory function of natural language syntax may be at issue,[2] but also the way in which individual items in the mental lexicon chunk otherwise unwieldy amounts of data into discrete, quickly accessible and manipulable packages may also be a key factor in gaining our uniquely human cognitive edge. This point was made by A. H. Bloom (1981, p. 76), and was later developed by me into a theoretical construct termed the *lingpack* (linguistically packaged molar concept) in my tentative model of the language-thought relationship (Henser 2000b), a model that embodies Levinson's insights into a possible way in which the "psychic unity of mankind" can be reconciled with conceptual differences across groups of natural language speakers (Levinson 1997, pp. 27–28). Were Carruthers to incorporate these insights into his theoretical framework, it would increase its scope and make it an even more persuasive argument for scientifically investigating the cognitive functions of natural language.

Work in the field of bilingual cognition has already produced interesting evidence of the inner deployment of language-specific resources and the cognitive difference that such resources make. The Spelke and Tsivkin study cited by Carruthers (Spelke & Tsivkin 2001) is one such example, showing that a specific NL is involved in the internal representation of large, exact numbers in mathematical calculations. Also, certain memories have been shown to be encoded in the language in which the relevant experience took place, irrespective of whether or not this context was the subjects' first language (Schrauf & Rubin 1998; 2000). Bilingual psychotic patients have been reported to lose their linguistic competence or have hallucinations in only one of their languages (Hemphil 1971; Laski & Teleporos 1977).

Widespread anecdotal evidence from bilinguals (cf. Grosjean 1982; Wierzbicka 1985) suggests that bilinguals feel and behave differently when using their different languages, and that there is a sensation of switching mental channels when switching from one language to another, either overtly or covertly. Some small-scale studies have been done into the causes of these reports (cf. Henser 2000a), but not since the extensive work of Ervin-Tripp back in the 1960s (Ervin-Tripp 1961; 1964a; 1964b) has there been any concerted attempt to explain these phenomena. Such questions cry out, both for thorough empirical research, and for a coherent and unifying theory to account for the role of natural language in our mental lives. Abandoning needless aversion to relativism may be an important step forward in making progress towards resolving these issues and gaining a fuller understanding of the nature of the language-thought relationship.

ACKNOWLEDGMENT
I would like to thank Andy Clark (Philosophy and Cognitive Science Programme at the University of Indiana, Bloomington, IN) for his helpful comments.

NOTES

**1.** Although these high-level representations are said to be in logical form (LF) (cf. Chomsky 1995), Carruthers makes it clear that this is a *natural language-specific* LF (see Carruthers 1996, pp. 249–50; Carruthers 1998, pp. 100–101).

**2.** Jackendoff has recently argued powerfully against a syntactocentric conception of language (Jackendoff 2002).

## Viewing cognitive mechanisms in the context of biology

Linda Hermer-Vazquez

*Department of Physiology, SUNY Downstate Medical Center, Brooklyn, NY 11203.* **Linda.Hermer-Vazquez@downstate.edu**

**Abstract:** Cognitive mechanisms are based in organisms' biology, and results from biological studies suggest that there is unlikely to be a single mechanism for reorienting or for combining information across modules or domains. Rather, there are likely to be multiple, partly overlapping systems for accomplishing nearly all cognitive and behavioral goals, as is the case for biological mechanisms more generally.

I agree overall with Carruthers's thesis that language provides a medium in humans for combining information across modules (with the loose definition of *module* that he uses) to form novel representations much more rapidly than would occur otherwise (for a review, see the discussion in Hermer-Vazquez et al. 2001, and for the original proposal, see Hermer 1993). Moreover, I like his suggestions concerning the role for inner speech in mediating the formation of these new representations. It was the fact that adult subjects commonly described using inner speech to conjoin geometric and nongeometric information in our earliest reorientation studies (Hermer & Spelke 1994) that led me to focus on the possible cross-module conjunctive powers of language, although of course these self-reports may have been describing an epiphenomenon rather than the actual conjunctive mechanism. In subsequent discussions, Elizabeth Spelke and I developed a hypothesis for how language comes to play this role in children learning language, relying on the possibility that the argument structure of language may not take into account domain- or module-related limitations. If this hypothesis is correct, then when children have learned terms for colors and sense (right and left) relations, the two can be combined into prepositional phrases such as "to the left of the blue wall" because the rules for language specify, for example, that any adjective – *blue* (nongeometric), as well as *long* (geometric) – can modify the final noun in a prepositional phrase that is otherwise about geometric (sense) relationships. This possibility as well as the ones that Carruthers mentions await further investigation.

All but the most dualist among us believe that such cognitive mechanisms are underpinned by biology, and as someone who has been conducting research in the fields of neurophysiology and behavioral neuroscience for the past six years, I would like to focus the rest of my comments on what biology suggests to us about the nature of cognitive modules and learning new representations. First, language is not likely to be *the* mechanism for the formation of cross-module representations, but rather, *one* mechanism among many, albeit one with particular advantages such as high speed (Hermer-Vazquez et al. 2001). Likewise, the geometric reorientation process and the process for locating objects with respect to movable landmarks (not discussed in the target article but a focus of Biegler & Morris 1993; 1996; and Hermer-Vazquez et al. 2001), are again two mechanisms among many with partly overlapping functionality. It is becoming increasingly clear that living organisms possess multiple, partly redundant, and partly unique coding and control mechanisms, all the way from the genetic level (Lobe & Nagy 1998; White 2002) to the level of large-scale neural circuits (Canedo 1997; McDonald & White 1993). In support of this for the cognitive level, young children conjoin geometric and nongeometric information to reorient themselves in larger environments (Learmonth et al. 2001; in press), and rats eventually learn to locate an object in the correct spatial relationship to a movable landmark (Biegler & Morris 1996). These findings are consistent with the fact that there is a hierarchy of cues used for navigation and object search more generally (Maaswinkel & Whishaw 1999), and that information is used to solve a given task appears to depend on such factors as how aversive the task is in addition to the task's cognitive demands (Gibson et al. 2002). Spanning over four billion years of the evolution of life, these partly redundant mechanisms have come to make for robust organisms that are not completely felled by the incapacitation of one system (e.g., Pettersson et al. 2000). These partly overlapping systems are also beneficial because they appear to "compete" to provide the best representation of a solution to a given problem (Poldrack et al. 2001).

Another important point from biology is that cognitive psychologists ought to avoid placing their findings about human abilities in a *scala naturae* framework. Recently, rhesus monkeys were found to use nongeometric as well as geometric cues for reorientation (Gouteux et al. 2001), as were a species of fish (Sovrano et al. 2002). Even the fish conjoined these two types of information together to solve the reorientation task given them. As Sovrano et al. point out, these findings indicate that a given species' abilities are better viewed in the context of adaptations to their particular ecological niche rather than in terms of that species' evolutionary distance from humans.

Finally, all these findings from biology and behavioral neuroscience underscore the difficulty of defining what a module or a domain is. For example, we originally defined the geometric reorientation process as a module because of its apparent encapsulation in adult rats and young children to all but information about the macroscopic shape of the environment (Cheng 1986; Gallistel 1990; Hermer & Spelke 1994; 1996). But when later studies showed that children conjoin geometric and nongeometric information in larger testing rooms, should it lead us to posit that there are two (or more) reorientation modules, taking in different types of information depending on room size? Even if we were to think that was a good inference, it would be circular because there would be no independent determination of module status. With multiple, partly overlapping processes representing problems and their solutions in parallel, more sophisticated conceptual frameworks are needed. Cognitive psychologists, with their elegant studies of operations on mental representations; biologists, with their techniques to study the full complexity of the systems in question; and philosophers, with their sharp argumentation, should work together to develop new ways of thinking about the robustness and messiness of living systems.

## The problematic transition from specific competences to general competence

James R. Hurford[a] and Jean-Louis Dessalles[b]

[a]*School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, Scotland, UK;* [b]*Groupe : Information, Interaction, Intelligence : I3 Ecole Normale Supérieure des Télécommunications, Paris, France.* **jim@ling.ed.ac.uk** **http://www.ling.ed.ac.uk/~jim/** **jld@inf.enst.fr** **http://www.enst.fr/~jld**

**Abstract:** Postulating a variety of mutually isolated thought domains for prelinguistic creatures is both unparsimonious and implausible, requiring unexplained parallel evolution of each separate module. Furthermore, the proposal that domain-general concepts are not accessible without prior exposure to phonetically realized human language utterances cannot be implemented by any concept-acquisition mechanism.

Carruthers draws heavily on Mithen (1996) for his argument that the thoughts of prehumans were sealed in separate compartments such as social understanding (for dealing with the conspecifics in one's group), technical understanding (for dealing with tools), and understanding of the natural world (for foraging and hunting). Mithen's thought-provoking account is largely arrived at by speculation from the archaeological record, rather than by experimentation on nonhumans. Not much about detailed brain organization can be gleaned from remains of bones, tools, and campsites. We are skeptical about this proposed modularity, on the grounds that it is unparsimonious and raises more problems than it solves.

What Carruthers calls conceptual modules (which include geometry, colour, and theory of mind) are claimed to produce nonconscious representations, like THE TOY IS IN THE CORNER WITH A LONG WALL ON THE LEFT AND A SHORT WALL ON THE RIGHT. This is, according to Carruthers, what animals and prelinguistic children have in their minds. This kind of recursive and syntactically complex form of mentalese is, according to Carruthers, module-specific. All these complex forms of representation are sitting there, waiting for the evolution of the language faculty to connect them: "We are supposing that this interface became modified during the evolution of language so that thoughts deriving from distinct conceptual modules could be combined into a single natural language sentence" (sect. 6.2, para. 2).

The assumption of the existence of $k$ different and yet compatible mentaleses is nonparsimonious. If all these representational modules evolved independently, we would need to explain how they arrived at forms compatible enough to be simply connected by the evolution of one further domain-general module. Similar mechanisms would need to be found for all ($k - 1$) other modules (colour, movement, audition, kinaesthesia, theory of mind, etc.). Carruthers needs $k$ parallel stories, with no guarantee that they will converge on compatible forms. Thanks to some further miracle, these mentaleses would be compatible enough to be easily combined in linguistic sentences.

For example, if predicate-argument structure is common to all specific modules, then to the extent that these modules are truly independent, separate explanations would have to be given for the emergence of this common structure. But if similar conditions each domain give rise to predicate argument structure (e.g., the existence of middle-sized salient objects, to which attention can be drawn, as suggested by Hurford 2002), then the domains are not completely distinct. It is surely implausible that there is no overlap between mechanisms for thinking about social life and hunting or foraging. Chimpanzees, for example, hunt in social groups for colobus monkeys. One chimpanzee will block the prey's exit in one direction, while another approaches from the other direction. This involves integration of social cooperation with hunting behaviour. At another level, the mechanisms involved in seeking and grasping nits during social grooming are likely to be similar to those involved in seeking and grasping insects for food, from the ground; grooming seems to be a case of hunting inside social interaction. Can we suppose that a chimpanzee's thoughtful strategies for catching nits during grooming are isolated from its strategies for catching insects from the forest floor?

Another separate difficulty with Carruthers's proposals involves his view that language can be a necessary condition for certain kinds of thought, yet not constitutively implicated in those forms of thought. Carruthers explains this initially perplexing view with a diachronic/ontogenetic proposal. The picture is that, during childhood, a person hears a full-fledged utterance with some domain-general content, interprets this by accessing its LF, and thereby gains access to a kind of thought previously inaccessible to him/her. The child's first access to such a thought *must* be through (fully fledged) language, but in later life, the person can access this thought directly via its LF, and the phonological trappings do not necessarily come to mind. This is Carruthers's position.

The interpretation of any utterance is seldom, if ever, accomplished in purely bottom-up fashion. That is, some contextual clues as to the general kind of meaning intended by the speaker are always deployed by the hearer in accessing the meaning of an utterance. Carruthers's account implies that children interpret utterances with domain-general content, which open up certain classes of thought to them, in purely bottom-up fashion, with no prior inkling of the LF treasure they are going to find at the end of the interpretation process. But this seems most implausible. Children, in particular, who know less of their language than adults, must rely *more* on top-down clues to meaning than do adults.

We need to ask how a child first learns the meaning of a particular domain-general natural-language predicate word, such as English *same* or *individual* or *three.* According to Carruthers, the child must hear an utterance using the word, for example, "This nut is the same colour as my skin," and without having prior access to the concept SAME, interpret the utterance correctly, and thereby gain access to the concept SAME. We are in a Fodorian paradox here, and Carruthers doesn't show us how to get out of it. If "non-domain-specific, cross-modular, propositional thought depends on natural language," there is an ontogenetic bootstrapping problem. Nonspecific thoughts must rely on domain-general concepts like SAME, INDIVIDUAL, or THREE. No account is given of how such concepts are acquired. Furthermore, Carruthers does not tell us where such concepts are supposed to reside because they do not belong to any specific module. If they are ascribed to a "central" system, then a $(k + 1)$th mentalese has to be postulated. Such a central system could, both ontogenetically and phylogenetically, play the role of integrator. If so, there is no need to grant this role to syntax and phonology. If there is no such central conceptual system, then the existence of any domain-general concept is problematic. Words alone, and their juxtaposition through syntax, can't bring them to existence in the mind.

ACKNOWLEDGMENTS
We thank the Caledonian Research Fund and the Royal Society of Edinburgh for a Visiting Research Fellowship for J.-L. Dessalles, which made this collaboration possible.

## How perspective shift integrates thought

Brian MacWhinney
*Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213.*
**macw@cmu.edu** **http://psyling.psy.cmu.edu/brian/**

**Abstract:** Within the context of Carruthers's general analysis of the relation between language and thought, I present a specific hypothesis about how grammar uses perspective flow to unify disparate cognitions. This perspective hypothesis allows us to understand the neolithic burst in creativity as a cultural advance in methods for knitting together thoughts.

The task of characterizing the relation between language and thought is a core challenge facing cognitive science. Carruthers's cogent analysis of this difficult problem is a welcome contribution to the literature. It was a particular pleasure for me to find that the specific shape of the language-thought relation Carruthers envisioned was so close in many ways to that which I have also been developing (MacWhinney 1977; 1999; 2003; in press). I have been arguing that language uses a system of perspective-switching to knit together thoughts from at least four cognitive levels: direct perception, spatio-temporal navigation, causal action, and social positioning. Like Carruthers, I believe this system builds up temporary mental models that then guide further thought. I also recognize that the constituent structure of syntax facilitates integration across these systems. Like Carruthers, I recognize that the cognitive powers of prelinguistic children, aphasics, and nonhuman primates show that we cannot characterize all thought as inner speech.

Although our views seem to occupy the same space in Carruthers's top-level logical analysis, they make extremely different claims about (1) the dynamics of the integration across systems, (2) the shape of the arena where integration occurs, (3) the nature of the evolution of the linkage between language and thought, and (4) the modular quality of the human mind. Let me, in this brief space, present an abstract of my alternative vision, which I call the Perspective Hypothesis.

My interest in the perspective hypothesis began when I realized (MacWhinney 1977) that the linguistic devices of initial positioning and subject marking were reflexes of "starting points" for cognition. Embarrassingly, it took me 20 more years to understand the extent to which this system of perspective flow permeated all of the constructions of language and not just the system of subject marking. In the meantime, additional experimental work, Gernsbacher (1990) and growing evidence for the incremental nature of sentence processing, made me confident that the notion of perspective maintenance and switching was core to understanding the relation between language and thought. The basic idea is that sentence comprehension involves an attentional flow that constructs the actions of a protagonist engaged in activities on the four levels of the theory: direct perception, space-time, causal action, and social relations. For example, in the sentence, "The dog the policeman was chasing jumped over the wall," we begin interpretation by assuming the perspective of the dog. We then shift to the policeman to imagine the chasing and then back to the dog to envision the jumping over the wall. Each of these perspectives is stored in verbal working memory, as we build up the dynamic image of enactive representations. In effect, this process reduces the whole of perception to the operations of a central homunculus that represents the self. This construction occurs immediately and incrementally as we listen to sentences. In production, our sentences trace out a set of activations using constructions and constituent structure in a way entirely consistent with Carruthers's account of the union of color adjectives with directions and place indexicals in a phrase such as "the short blue wall."

Unlike Carruthers, I view this high-level homunculus as more than just a call-back to input modules. For me, central cognition occurs not through a feedback loop to the input modules, but in an arena constructed by the interaction of the various linked posterior-anterior projections areas found in frontal cortex. The evolutionary construction of this frontal homunculus as a device for organizing planning began in the primates and continued across the full six million years of hominid evolution. While agreeing with Mithen's arguments about the separateness of cognitive systems in our ancestors, I would choose to follow Donald (1998) when he argues that the beginning of an integration across these systems in *Homo erectus* well antedates the emergence of spoken communication. I interpret the neolithic sociocognitive breakthrough studied by Mithen as the result of cumulative changes in temperament and cognition, which then interacted with ongoing advances in vocal capabilities to achieve the solidification of new social institutions, such as the priesthood. The fact that radical social changes during the last 20,000 years have led to an integration of the human mind shows that language has achieved its effect not through some remarkable recent mutation, but by making new use of systems that we already in place.

What does this analysis say about modularity? First, it is clear that the kinder, gentler, emergent modules proposed by Mithen, Carruthers, Tooby, Cosmides, Pinker, and Spelke bear little resemblance to the modules proposed originally by Fodor. Beyond this, one wonders whether it might not be better to think of the effects of perspective flow not in terms of breaking down the walls between modules, but in terms of facilitating communication between distant areas of the brain. By emphasizing the ability of language to knit together discrepant perspectives over cognitive time, we avoid a questionable enumeration of specific modules and a largely fruitless debate regarding relative permeability and biological determination. The perspective hypothesis allows us to understand the extent to which language "runs all around the brain" grabbing information from whatever source it can find. This ability to integrate information is not contingent on the exact modular shape of the informational loci. Instead, it depends on the integrative facilities of language. There are three ways in which language achieves this integration neuronally. First, lexical items receive stable input projections from a variety of cognitive inputs (Li et al., in press) Second, as Carruthers notes, constituent structure allows for "plug-ins" through modification and attachment. Third, by establishing a central perspective to each predication and tracking shifts in that perspective, connected discourse can establish a full social narrative. The cumulative effect of these three integrative processes is to draw on all aspects of cognition to construct a central cognitive representation grounded on a fictive, intentional homunculus.

## Could you *think* Carruthers's ideas without having to *speak* them? *Talk* with yourself if you want to have any *thought* on that

Carlos Molina

*Instituto de Filosofía, Pontificia Universidad Católica de Chile, Santiago, Chile.* **cmolinac@puc.cl**

**Abstract:** First, the importance of language in cognition is recognized. Nevertheless, this does not necessarily imply that the *locus* of thought is natural language (words, syntax, phonology). Then, difficulties with some of Carruthers's hypotheses are stated: Is an account based on LFs capable of dealing with the complexities involved in what we call *thought?* Finally, mention of the issue of language production is made.

Carruthers puts forward the hypothesis that "there are *some* thought-types which can only be entertained at all, by us, when tokened consciously in the form of an imaged natural language sentence." This observation is basically obtained via introspection of our own thought processes. It's quite plausible to think that the architecture of human cognition has this property. We have the impression that if someone asks, "say, what do you think about Carruthers's ideas?", you can think about the issue only if you *aid* your thinking with natural-language sentences. If the following day you meet that person and, for some reason, you didn't talk to yourself about the issue in the meantime, it is probable that you wouldn't have anything to say. Because you didn't *talk* about the issue, you didn't *think* about it.

It is important to add that this does not mean that the conceptual system is unable to think (make inferences, etc.) if the linguistic representation levels (roughly words, syntax, and phonology) are not activated. In fact, surely there is a lot that non-linguistic-aided thinking (NLAT) can do. We would like to point out that *some* thought processes (roughly, complicated reasoning) require the aid of the language faculty resources. But this doesn't necessarily mean that the *locus* of thought is natural language representations (words, syntax, phonology). This is what Carruthers calls the *supra-communicative* conception of language, and it is the one that we endorse.

Let's now turn to some of Carruthers's arguments. Remember that one of Carruthers's hypotheses is that knowing a natural language such as English or Chinese is essential for complicated reasoning because such reasoning requires *access* to the sort of representational system that can only be provided through learning the structure of a natural language. But, what does this representational system consist of? Carruthers proposes Chomsky's characterization of language and hypothesizes that logical forms (LFs) are the level of representation that constitute the *locus* of thought (where thinking is really happening). In a second hypothesis he claims that LFs are the mechanisms that allow the connection between different conceptual domain-specific modules (e.g., folk physics, folk biology, ToM, etc.). We find it strange that he doesn't take greater care in characterizing with precision the notions of *language* and *thought.*

The notion of language for Carruthers is clearly based on what is meant by it in Chomsky's theories (Chomsky 1995).[1] Specifically, he believes that Chomsky's LFs are the basic building blocks of thought. But why is this? It remains unexplained, but it seems to be that the syntactical properties of LFs are the suitable medium for explaining the fact that we can entertain thoughts such as, "The food is in the corner with the blue wall on the left," where we connect "blue' (color module) with "wall" (geometrical module). In this respect he says, "natural language syntax allows for multiple embedding of adjectives and phrases . . . so there are already 'slots' into which additional adjectives – such as 'blue' – can be inserted." But this doesn't look like an explanation! And he doesn't say much more about this matter in the article. Could it be the case that the author believes it is not really necessary to make those mechanisms more explicit? Aren't those mechanisms at the heart of what is at stake?[2] I have to confess that I don't know exactly what the *locus* of thought is (who knows?),[3] but whatever it is, it will have something to do with the way we organize experience; and this is clearly not in terms of adjectives, nouns, and verbs, but in terms of "conceptualised entities" such as objects, events, properties, times, quantities, relations, intentions, and the like. And as linguists have shown us, the mapping from conceptual category to syntactic category is many-to-many (see Jackendoff 1983; 2002).[4]

In addition, Carruthers points out that "the present hypothesis commits us to a nonclassical account of speech production." But, is Carruthers's hypothesis a genuine alternative to Levelt's (1989)?[5] In Carruthers's view, there is no such entity as a "preverbal thought" because a thought "cannot be entertained independently of being framed in natural language" (sect. 6.1, para. 1).[6] OK, but why? Again, no explanation is given. It looks very much like an ideological statement. Carruthers remarks that neither he nor Levelt has an answer to the question of how thoughts get assembled in the first place. In a certain way this is true, but there are some important differences between both views: In Levelt's account, the process of generating messages is analysed as a two-step process, with macroplanning followed by microplanning. Roughly speaking, macroplanning has to do with deciding what to say and in what order; microplanning has to do with giving the units of macroplanning an information structure, a propositional format (as opposed to, say, a visual format), and a perspective. The final result of microplanning is a "preverbal message" that can be recognized by the language faculty as its characteristic input (see Levelt 1989). All these processes are not capriciously stated but are the result of taking into account the distinctions that speakers would have to make to produce language. How could LFs deal with all these different types of processes? This also remains unexplained. Some other problems arise, such as how to deal (in terms of LF) with the reasonable idea that, "there has to be a 'what we meant to say' that is different from what we said" (Pinker 1994, p. 57).

In sum, we think that Carruthers makes a valuable contribution by making us take into account an important phenomenon: When we engage in complicated reasoning, the language faculty is required. So far so good, but the problem arises when the author establishes hypotheses concerning the mechanisms behind this phenomenon. We believe that the mechanisms he puts forward (LFs, etc.) are too general and unspecific. An adequate understanding of the relation between language and thought has to point to specific features of both language and cognition. Without clear and adequate operational definitions, we cannot investigate general assertions about relationships between language and thought.

NOTES

**1.** Chomsky identifies the level of language that interfaces with the conceptual-intentional system as *logical form* (LF), a level built out of syntactic units such as noun phrases and verbal phrases. Remember that Chomsky characterizes generative grammar not as a method for constructing sentences but as a formal way to describe the infinite set of possible sentences (competence). The fundamental generative component of the system is the syntactic component; the phonological and semantic components are "interpretive." It is curious to find that Carruthers was inspired by a theory that had no semantic motivations, and that is not concerned with linguistic performance, all this given that his main concern is to understand "how we think."

**2.** Remember that in Carruthers's own views the problem is a natural one, not a conceptual one.

**3.** "It is, of course, exceptionally difficult to determine how people 'really' represent situations to themselves" (Slobin, in press).

**4.** "It is widely accepted that syntactic categories do not correspond one to one with conceptual categories. All physical object concepts are expressed by nouns, but not all nouns express physical object concepts (consider *earthquake, concert, place, redness, laughter, justice*). All verbs express event or state concepts, but not all event or state concepts are expressed by verbs (*earthquake* and *concert* again). Prepositional phrases can express places (*in the cup*), times (*in an hour*), or properties (*in the pink*). Adverbs can express manners (*quickly*), attitudes (*fortunately*), or modalities (*probably*)" (Jackendoff 1997, p. 33).

**5.** The most sophisticated theory on the market and the result of years of empirical and theoretical research.

**6.** It is assumed that Carruthers is referring here to structured or complex thoughts, such as THE TOY IS TO THE LEFT OF THE BLUE WALL.

## Developing dual-representation processes

Katherine Nelson

*Department of Psychology, CUNY Graduate Center, New York, NY 10016.*
**knelson@gc.cuny.edu**

**Abstract:** Cross-domain representations provide the foundation for language and are not its unique product. Modularity of a limited kind is confined to early infancy in humans and is succeeded by domain-general thinking and speaking. Representational language becomes accessible to the cognitive system during the preschool years as a supplement to experientially based conceptual processing, resulting in a dual-process system.

Carruthers asserts that "*all* non-domain-specific reasoning of a non-practical sort . . . is conducted in language" (sect. 5.1, para. C) because the nonlinguistic conceptual system is characterized by widespread domain-specific modularity that inhibits many kinds of conscious intelligent inference and reasoning. From a social-cognitive-developmental perspective, none of this necessarily follows. Nonetheless, there is a place for a dual process theory in cognitive development.

Consider first the evolution of language as a communicative system, certainly its primary function. Although its beginnings are controversial and obscure, it can be said with confidence that it emerged as a social symbolic system constructed from combinations of symbols whereby representations of the external world as well as internal thoughts could be shared among groups of humans. Shared linguistic representations provide the basis for collective reasoning, as in the debates among hunters that Carruthers cites. It seems plausible that such collective cognitive uses of language were the platform from which individual thinkers derived language representations and associated linguistic based reasoning as a supplement to the basic cognitive conceptual system, resulting in a dual process cognitive architecture. Such a sequence gains plausibility from similar social to individual sequences in development, learning, and history (Scribner 1985; Vygotsky 1986).

But is it plausible that individuals participating in social colloquies did not already possess nonlinguistic, non-domain-specific conceptual processed? If they did not, how would they comprehend or interpret the cross-domain discourse of others? Or, how would any of the group be able to compose linguistic representations that crossed domain boundaries? Similar questions must be raised with respect to development. How could language – a non-domain-specific symbolic communication system – be acquired by infants or children if they did not already possess a non-domain-specific conceptual cognitive capacity? Language is not in fact acquired in a domain-specific mode and then somehow transformed into a nonspecific system. Even at the first one-word level, infants are not confined to learning words within one ontological domain such as objects, but range over a variety of conceptual domains such as persons, places, times, natural phenomena, events, actions, social phrases, and so on (Nelson et al. 1994). And exchanging the simplest utterances requires combining the perspective of speaker or listener (intuitive psychology) with a message that may refer to physical objects, actions, events, possessions, and so on.

The point is that language does have the potential for combining information from different conceptual domains, but that to acquire language, infants and children must have a conceptual platform adequate to the task, namely, one that does not constrain language use to a single domain. (The earlier claim that language learners focused solely on words for objects has been abandoned in the face of variability of vocabularies, individual differences in types of words learned, and cross-linguistic variability.) Two questions follow: What status then do domain-specific modules have? What sort of conceptual representation supports nonspecific language learning? A different sort of theory of early cognitive development than that assumed by Carruthers and other cognitive scientists in his camp is needed.

The importance of modularity in human thought may be questioned without doubting either that the intelligence of other animals, including our close primate relatives, may be largely domain-specific, sculpted by natural selection; or that human infants are born with predispositions to process certain classes of perceptual stimuli (faces, objects, phonology, speech patterns, spatial relations) in modular ways. A plausible cognitive developmental sequence recognizes the adaptive significance of such processes to early infant life and the laying down of basic neural structures. These processes engage experience-expectant neural systems and establish basic parameters regarding the environment, including its critical social components. Beginning in the middle of the first year, the infant cognitive system becomes less dependent on modular functions and more plastic, more experience-dependent and domain-general, setting the stage for socially shared attention, event representations, imitation, and language acquisition. That event representations serve as basic cognitive structures beginning at least in the last half of the first year is relevant to the claim that language acquisition itself rests on domain-general processes (Nelson 1986). Representations of events (e.g., eating and going to bed) involve cross-domain integration of persons, objects, actions, space, and time, and there is good evidence that such representations provide conceptual input to early language (Toma-sello 1992). Delayed imitation in later infancy is based on event representations and provides a prelinguistic social-experiential semi-symbolic cognitive representation (Donald 1991). At the perceptual, motor, and nonsymbolic level of the cognitive system, automatic nonconscious processes continue to subserve both the nonlinguistic conceptual and linguistic representations. Non-domain-specific conceptual processes, including language, may be thought of as layers added to earlier evolved cognitive operations.

A transition period of language practice in social contexts continues up to the fifth or sixth year when linguistic representations and processing supplement the basic experience-based system. I have proposed a dual process theory to explain the important advances in cognitive development during the preschool years (ages 3–5) (Nelson 1996; in press; Nelson et al. 2003). Specifically, the achievement of *representational language* supplements the initial nonlinguistic experientially based conceptual system with a second linguistic representation component available to the cognitive processing system. Among other things, this enables holding in mind two competing or complementary representations at the same time (e.g., past and present states of affairs, first- and second-person perspective on the same event, etc.), with implications for performance on theory of mind tasks and category choice tasks, as well as self-reflective thought. The achievement of representational language involves the acquisition of complex semantics and syntax necessary to express propositional thoughts about things not evident in the immediate scene, such as descriptions, explanations, narratives, plans, and so on. This level of language competence, which follows several years of conversational practice with adults, particularly those concerned with the past and future, is reached by about four to five years of age.

Subsequently, domain-specific conceptual structures and associated learning and memory strategies begin to be constructed and consolidated, in collaboration with the social-cultural world that defines useful domains such as writing and arithmetic, poetry and physics, or alternatively, hunting and gathering, spinning and woodworking (see also Karmiloff-Smith 1992). Such domain-specific knowledge does not of course prevent children or normal adults from engaging in cross-domain thinking, either in science or in everyday life.

In summary, language is a cross-domain communicative and cognitive representational system made possible by a preexisting non-domain-specific conceptual system perhaps unique to human cognition. Together they compose a dual representation cognitive processing system of the extraordinary flexibility and power that we are familiar with.

## Internalizing communication

Gerard O'Brien and Jon Opie

*Department of Philosophy, University of Adelaide, Adelaide, South Australia 5005.* **gerard.obrien@adelaide.edu.au**
**http://www.arts.adelaide.edu.au/philosophy/gobrien.htm**
**jon.opie@adelaide.edu.au**
**http://www.arts.adelaide.edu.au/philosophy/jopie.htm**

**Abstract:** Carruthers presents evidence concerning the cross-modular integration of information in human subjects which appears to support the "cognitive conception of language." According to this conception, language is not just a means of communication, but also a representational medium of thought. However, Carruthers overlooks the possibility that language, in both its communicative and cognitive roles, is a nonrepresentational system of conventional signals – that words are not a medium *we think in,* but a tool *we think with.* The evidence he cites is equivocal when it comes to choosing between the cognitive conception and this radical communicative conception of language.

Reflection on our phenomenology furnishes two very different views of the role of natural language in human cognition. On the one hand, as Carruthers observes at the beginning of the target article, we are constantly running words and sentences through our heads, even when performing quite trivial cognitive tasks. This provides some support for what Carruthers terms the "cognitive conception of language," according to which natural language constitutes a representational medium of thought. On the other hand, there is the familiar feeling that our thoughts are present in some form before we attempt to express them in natural language ("I know *what* I want to say, I just don't know *how* to say it"). This feeling is more consistent with what Carruthers terms the "communicative conception of language," which denies that natural language is a representational medium of thought and holds instead that its role is restricted to the communication of thoughts coded in other representational media.

Because phenomenology doesn't settle this question, and because one of his aims is to defend the cognitive conception of language, Carruthers turns to the empirical evidence. His trump card is a series of experiments conducted by Hermer-Vazquez et al. (1999), which suggest that natural language has a crucial role in the human capacity to integrate information across different domains (in this case geometric information and information concerning object-properties). Carruthers concludes (as do Hermer-Vazquez et al.) that natural language performs this integrating function by acting as a mental *lingua franca:* a medium for representing items of information drawn from distinct domain-specific modules in the brain.

In our view, however, the situation is not so straightforward. We accept, on the basis of the experimental evidence Carruthers presents, that natural language contributes to the cross-modular integration of information in human subjects. What we question is the further claim that it does so *by acting as a representational medium* (and hence that the cited experimental work vindicates the cognitive conception of language). Carruthers does allow that language can influence cognition *without* acting as an (internal) representational medium, because it both sculpts and scaffolds cognition (sect. 2). But, as he points out, the function of language in these cases is primarily diachronic – it contributes to the *acquisition* of certain cognitive capacities rather than the real-time *exercise* of those capacities. Consequently, such processes are incapable of accounting for the synchronic integration of internally coded information, and the cognitive conception of language appears to be obligatory. However, we believe that Carruthers has overlooked a significant cognitive role for language that is both synchronic and consistent with the communicative conception – one, moreover, that has the potential to account for the cross-modular integration of information in human subjects. To demonstrate this we'll need to consider the communicative conception of language in more detail.

What distinguishes the communicative and cognitive conceptions of language is that the former excludes natural language from among the human brain's representational media of thought. Natural language is in the communication business, not the thinking business. Most proponents of this view accept that language enables communication precisely because it is a representational medium, but deny that it is a medium of thought. Some go further. Robert Cummins, for example, argues that written and verbal tokens of natural language communicate thoughts, not by representing them, but merely by acting as conventional "signals" that trigger appropriate representing vehicles in target brains (see especially Cummins 1996, pp. 135–40). Such signals are produced when representing vehicles in a source brain interact with motor systems via mechanisms that realize the governing conventions of language. And they influence the receiving brain by impacting on its sensory surfaces. Communication succeeds when the emitted signals cause the receiving brain to token vehicles in its representational medium of thought whose contents are sufficiently similar to those tokened in the source brain. Cummins's insight is that linguistic signals need not (indeed *should* not) be conceived as content-bearers to explain their role in this process.

This more radical communicative conception of natural language is echoed in a recent paper by Paul Churchland, who adds an interesting twist. He claims that language is not a system for representing the world, but the "acquired skill of perceiving . . . and manipulating . . . the brain activities of your conspecifics, and of being competent, in turn, to be the subject of reciprocal brain-manipulation" (Churchland 2001). From this perspective, language is not just a means of communicating thoughts but also a way for one brain to shape the cognitive activities of another – linguistic signals fundamentally operate as a means by which we manipulate the contents and trajectory of thought in other people.

What sort of cognitive role does this leave for natural language? Clearly, sculpting and scaffolding are not ruled out, but a far more significant possibility is that the communicative process we have described also goes on *inside individual brains.* It was Vygotsky's great insight that after children acquire a natural language as a tool for communication, they internalise it, that is, they appropriate it as a cognitive tool (Vygotsky 1962/1986). For Vygotsky, as for Carruthers, this process is one in which an external communicative scheme becomes an internalized *representational medium.* What we are suggesting, by contrast, is that the internalisation of natural language is a process whereby a conventionally governed set of communicative signals is put to work inside a single brain. Language becomes a system of signals apt not only for manipulating the brains of others, but also for recurrent *self*-manipulation. Such internalisation involves establishing communicative links (routed through the language centres) that can be employed by one part of the brain to steer the cognitive activities occurring in other parts of the brain.[1] In this picture, natural language is not an internal representational medium, it is a powerful cognitive tool – one that establishes coherent, multimodal representations, by facilitating communication within the brain, and regulates the sequencing of thought via the constant interplay between networks that encode linguistic signals and those that encode thoughts (for a fuller discussion, see O'Brien & Opie 2002).

There is emerging evidence that language, implemented primarily in temporal cortex, plays just this kind of role. Recurrent connections that loop from language centres out to the representing vehicles they trigger, and back again, catch up language and thought in a tight web of mutual influence that extends our cognitive capacities well beyond those of infraverbal organisms. Such connections may well function to integrate information across the brain's perceptual and conceptual faculties, without recourse to a linguiform medium of thought. What we are proposing is a synchronic, fully internalized cognitive role for language that is consistent with the communicative conception. As we have put it elsewhere (O'Brien & Opie 2002), words are not a medium we *think in,* they are a tool we *think with.* The possibility of this radical communicative conception of language is one that Carruthers has failed to consider. And the empirical evidence he cites is equivocal when it comes choosing between his favoured cognitive conception and the alternative we have sketched.

NOTE

**1.** This idea is somewhat similar to (but not identical with) the speculations that Dennett makes about the role of natural language in organizing our thinking (see, e.g., Dennett 1991, pp. 193–99).

## Speech as an opportunistic vehicle of thinking

Csaba Pléh

*Center for Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary.* **pleh@itm.bme.hu**
**http://www.itm.bme.hu/ktk/csaba**

**Abstract:** Carruthers clearly identifies the basic issues involved in language and thought relations and argues for an adaptive central model. Similar conclusions were reached by classical research in the inner speech tradition. Sokolov (1968; 1972) especially emphasized that inner speech appears only when the task is difficult. The use of inner speech is not a necessity to transform representations, but it is called for when transformations become difficult. This might be related to the cognitive reorganizations leading to language as emphasized by Donald (2001).

The clear differentiation by Carruthers between two basic questions involved in the complex language and thought interface issue, namely, the involvement of language as a system and speech as an actual bodily process, on the one hand, and the mandatory or adaptive involvement of linguistic systems in thought, on the other, is a most welcome effort in the domain of cognitive science. In my commentary I shall concentrate on two issues: What can we learn in this regard from traditional experimental work on dual tasks in the inner speech research tradition, and how should one envisage the shaping of domain general or cross-domain mechanisms in the genesis of human cognition?

As Carruthers points out, cognitive science usually starts off from a communicative conception of language when dealing with the language-thought relationship. This is, according to Carruthers, Fodor's (1983) input/output modules. There is, however, another work of Fodor (1975), the one Carruthers has in mind when he talks about the outputs of mental modules, that is, the Fodor of the language of thought, where an abstract proposition like representation is claimed to be the vehicle of any internal thought process. However, this claim is being made with no clear position regarding whether this *lingua mentis* is related to actual speech in either its genesis or its workings.

Determination in its different forms and scaffolding characterize the argumentation in the language and thought literature as Carruthers surveys it, from the conceptual analysis of philosophy through the studies on category use in cultural interactions down to laboratory research on inner speech. One can summarize and extend the different existing positions in Table 1.

Carruthers presents a rather well-argued view according to which central process modularity should be reconciled with a weakened version of linguistic determinism to the effect that language is used as a central mediator between modules, as the vehicle, if you wish, for the outputs of central processes – along the line of Fodor (1975) – which is used to compute complex trains of representations. I basically agree with this type of limited and adaptive linguistic determinism, and with the evolutionary arguments put forth in favor of it. One could enrich the evolutionary tracing by realizing how the scaffolding introduced by culture supplements or organizes the mental workings postulated by Carruthers. I have in mind specifically the position of Merlin Donald (1991; 1993; 2001) who claims that language and then writing involve feedbacks toward our own cognitive architectures (see Pléh 2002.) Donald claims, "We are a culturally bound species and live in a symbiosis with our collective creation. We seek culture, as birds seek the air. In return, culture shapes our minds, as a sculptor shapes clay" (Donald 2001, p. 300).

This is a nontrivial issue, because it touches on the way we conceive the relationships between the external social world, together with the communicative functions of language and the cognitive domain and functions. There is disagreement about this, but I do not wish to it solve here, just to allude to it with another quote from Donald (2001, p. 254). "The great divide in human evolution was not language, but the formation of cognitive communities in the

Table 1 (Pléh). *Different positions and research strategies regarding the language/thought interface*

| Who leads? | Lexicon and categories | Grammar and propositions | What research to do? |
|---|---|---|---|
| Language | Labels shape categories | Structured thought presupposes language (Wiitgenstein, Whorf, Carroll, Bowerman) | Category analysis; cross-language comparison of cognition |
| Speech | Naming facilitates recall of categories (Brown & Lenneberg) | Speech allows or facilitates long trains of thought | Inhibiting speech or dual tasks in cognition |
| Thought and mental representation | Universal cognitive theories (Piaget) | Logical structure is primary (Frege, Husserl) | Categorial analysis child development |
| Tied or unified | Jackendoff "common core" theory; bootstrapping models (Pinker, Gelman) | | Comparing computations in different domains |
| Independent at the outset but meeting later | Modularity Fodor and followers; Vygotsky culturalism; Karmiloff-Smith modularization | | Search for domain specificity; study of developmental interactions |

first place. Symbolic cognition could not spontaneously self-generate until those communities were a reality. This reverses the standard order of succession, placing cultural evolution first and language second."

My other point relates to Carruthers asking for "further evidence" where he clearly emphasizes the need for more studies using dual tasks. Interestingly enough, the dual task tradition has been with us in the domain of inner speech research, at times and frameworks when central process modularism has not yet been around. The initial doctrine in this domain was the behaviorist claim that inner speech is the material carrier of internal processes. However, we have the illusion that we think it is actually an activation of our speech muscles or the innervation of them that corresponds to our impression of thinking (Watson 1920; 1924). Partly because of the limitations of available technology, traditional inner-speech studies used the logic of dual tasks in proving this claim, looking for indirect evidence for the participation of speech in mental processes. One of the methods was to inhibit inner speech by asking the subjects, for example, to protrude their tongues while solving mental arithmetic tasks. This passive dual task was supplemented by active dual tasks where subjects had to involve their speech apparatus in an overlearned speech sequence, for example, reciting a well-memorized poem or trying to read silently during intellectual tasks (Sokolov 1968; 1972). McGuigan (1978) provides a detailed survey of these methods, labeling them distraction and interference studies. McGuigan, along the behaviorist tradition, tends to summarize the literature and his own studies by claiming that all mental processes require covert behavior, Sokolov (1968; 1972) arrives at a conclusion more in line with the adaptive ideas of present-day central models like Carruthers. There is no one-to-one correspondence between inner speech and thought. The active involvement of the speech apparatus in another task usually leads to decreased performance in mental arithmetic tasks, some visual memorization tasks, and, of course, many language tasks like translation. However, in easier tasks the rivaling articulatory dual task does not impede performance.

> This means that if thought acts remain stereotyped they can be performed in a more or less inhibited status of the speech movement system. Ordinary everyday experiences correspond to this experimental fact. It is well known, for example, that many thoughts come to our mind while we are engaged in conversation with others or we read something or lecture, i.e., when our speech apparatus is busy in pronouncing other words. (Sokolov 1968, p. 125)

Electrophyisiological studies also indicate that inner speech appears when the task becomes difficult, according to Sokolov, across different domains as well. Thus, in solving easy visual figural tasks from the Raven test, no inner speech is observed through electrophysiological measurements; it does appear, however, when the figural task becomes difficult. Functionally, as some of the debriefing reports imply, this would correspond to verbalizing the details of the representation and then comparing the verbal formulations like, "Well, there are two lines on that one, and three lines on the other one, these don't fit, but two here and two there. These belong together."

This claim is rather similar to conclusions coming from a different domain about the implied impact of linguistic labels on color recognition: The accessibility of labels (codability, as it was called by Brown & Lenneberg 1954) had an impact if the set was large or if the delay time was increased. This suggests that the task is performed in case of difficulties with the help of recoding and memorizing color names rather than being based on sensory experience itself.

## What domain integration could not be

Philip Robbins

*Department of Philosophy, Philosophy-Neuroscience-Psychology Program, Washington University in St. Louis, St. Louis, MO 63130.*
**probbins@artsci.wustl.edu**

**Abstract:** Carruthers argues that natural language is the medium of non-domain-specific thought in humans. The general idea is that a certain type of thinking is conducted in natural language. It's not exactly clear, however, what type of thinking this is. I suggest two different ways of interpreting Carruthers's thesis on this point and argue that neither of them squares well with central-process modularism.

According to central-process modularism, the bulk of postperceptual cognitive processing is accomplished by a plurality of special-purpose mechanisms, each dedicated to a particular task domain, such as mindreading, causal inference, or biological classification. Few if any advocates of this theory of cognitive architecture endorse the radical view that *all* central cognition is like this. Indeed, most are careful to leave space at the table for central processes that are not proprietary to any particular cognitive module, hence not domain-specific. No doubt this is a good thing (Fodor 2000). But at this point it becomes natural and important to ask: What features of our cognitive architecture explain those instances of cognition that transcend the modularist framework? Are these higher flights of thinking special to humans, and if so, what is it about us that underwrites this fact?

An increasingly popular view in contemporary cognitive science, especially among developmentalists, suggests that the key to answering these questions may be found in the domain of language. This is the view Carruthers seeks to promote. Buoyed by recent empirical findings of Elizabeth Spelke and her colleagues (Hermer & Spelke 1996; Hermer-Vazquez et al. 1999; Spelke & Tsivkin 2001a; 2001b), he argues that natural language is the medium of non-domain-specific thought in humans. The general idea is clear enough: A certain type of thinking, not proprietary to any particular central module, is conducted in natural language. What's less clear – indeed, what stands in need of clarification – is exactly what type of thinking this is. I will suggest two different ways of interpreting Carruthers's thesis on this point and argue that neither of them squares well with central-process modularism. The resulting dilemma threatens to undermine this particular variant of the cognitive conception of language.

The principal feature of the type of thought at issue is that it involves "integrating information" from distinct domains. But there are at least two ways to think about what such integration comes to. The first is that it is simply a matter of the representational content of a thought, where thoughts are structures figuring in certain computational processes (reasoning, problem solving, decision making, and the like). Thoughts have constituent structure, and their constituents – that is, concepts – may be drawn from one or another domain. On this view, then, a non-domain-specific thought is a thought that combines concepts from distinct cognitive domains. A paradigm case, drawn from Spelke et al.'s studies of spatial orientation, would be the thought (1) THE TOY IS TO THE LEFT OF THE BLUE WALL, which combines geometric concepts (LEFT OF) with color concepts (BLUE). Carruthers says that "our hypothesis is that such a thought cannot be entertained independently of being framed in natural language." The capacity for domain integration in this sense, then, is just the capacity to host thoughts built up from domain-heterogeneous components.

If this is how the notion of domain integration is cashed out, however, it becomes hard to believe that language is necessary for domain-integrative thought. To see why, recall the background thesis of central-process modularism. The idea is that most cognitive capacities are realized by a special-purpose conceptual mechanism: Mindreading is accomplished by a folk psychology module; reasoning about physical objects, by a folk physics module; spatial navigation, by a folk geometry module; and so on. But when we look into the representational contents of these various modules, we are sure to find representations at least some of whose components are not proprietary to the module in question. For example, Carruthers suggests that the geometrical module might construct the thought (2) THE TOY IS IN THE CORNER WITH A LONG WALL ON THE LEFT AND A SHORT WALL ON THE RIGHT, and the object-property system might build a thought like (3) THE TOY IS BY THE BLUE WALL.

Notice, however, that both thoughts implicate exogenous concepts (hence, exogenous information). The thought in (2) includes the nongeometric concepts TOY and WALL, and the thought in (3) includes both those concepts, neither of which is plausibly in the domain of object-properties, as well as the spatial concept BY. And examples like this can be multiplied. Consider the case of mindreading. A key component of mindreading is the capacity to attribute propositional attitudes to other agents. Such states are specified in two dimensions: by attitude type (belief, desire, intention, etc.) and by intentional content (*that snow is white, that lunch be served,* etc.). And representing the latter typically requires nonpsychological concepts (e.g., SNOW, WHITE, LUNCH). So here, too, domain integration in the weak sense just described seems inevitable even in cases of intradomain thinking. Because Carruthers insists that language is probably not necessary for this sort of cognition, the present notion of integration doesn't fit the bill.

Here is a second proposal. Let's say that a thought is domain-integrative just in case it combines information in the weak sense above *and* is an output of an inference whose inputs span two or more central modules. This stronger notion locates the property of informational integration at the level of processes rather than states. For example, the thought (4) THE TOY IS IN THE CORNER WITH A LONG BLUE WALL ON THE LEFT is domain-integrative – hence, non-domain-specific – because it is computed from (2) and (3). According to this view, the claim that non-domain-specific thought in humans depends on language, amounts to the idea that without language, certain kinds of transitions in thought would be impossible for us. Carruthers suggests that this is because the very process of making such transitions causally implicates the language faculty in an essential way. But how?

The answer on offer goes as follows (sect. 6.1). The geometry module and the object-property module, respectively, output (2) and (3), and both thoughts are taken up as inputs by the language faculty. Each is then translated into a natural-language expression, namely (5) [$_S$ The toy is in the corner with a long wall on the left and a short wall on the right] and (6) [$_S$ The toy is by the blue wall], and the results are somehow combined to yield (7) [$_S$ The toy is in the corner with a long blue wall on the left]. This is just a tentative sketch of how speech production might proceed in the case of non-domain-specific thought. But it does show how one might sidestep the standard supposition that thought invariably precedes its linguistic expression.

Unfortunately, the story just described hasn't much to recommend it beyond that. For it is doubtful that the language faculty could carry out the sorts of extralinguistic computations ascribed to it here. Speaking of the transition from (5) and (6) to (7), Carruthers notes that "it would not be too complex a matter for the language production system to take two sentences sharing a number of references like this, and combine them into one sentence by inserting adjectives from one into open adjective-slots in the other." But complexity is beside the point. The question is whether the language faculty itself should be thought to contain a general-purpose inference engine. And the likely answer – surely the answer one would expect from a central-process modularist, as well as the answer given by most linguists (Larson & Segal 1995) – is No. Language is one thing, reasoning is another.

## The spatial reorientation data do *not* support the thesis that language is the medium of cross-modular thought

Richard Samuels
*Philosophy Department, King's College, Strand, London, UK*
**richard.samuels@kcl.ac.uk**

**Abstract:** A central claim of the target article is that language is the medium of domain-general, cross-modular thought; and according to Carruthers, the main, direct evidence for this thesis comes from a series of fascinating studies on spatial reorientation. I argue that the these studies, in fact, provide us with no reason whatsoever to accept this cognitive conception of language.

Although there are many points of enormous interest in the target article, perhaps the *fundamental thesis* contained therein is that "language is the medium of [domain-general] cross-modular thought." And, according to Carruthers, the "most important" source of direct evidence for this thesis comes from a series of elegant studies on spatial reorientation (Hermer & Spelke 1996; Hermer-Vazquez et al. 1999; 2001). As he is well aware, however, these studies constitute evidence for the fundamental thesis *only if* they support the following *restricted thesis:* Language production processes – as opposed to central processes – are responsible for the integration of geometric and nongeometric information in the performance of reorientation tasks. In what follows, I argue that these studies do *not* support the restricted thesis, simply because they fail to provide us with any reason whatsoever to favour it over the competing, more traditional view that one or more *cen-*

*tral* system is responsible for the performance of such integrative tasks (Fodor 1983). My burden is thus to explain how this traditional view can account for the main results on spatial reorientation – the dual-task data and the correlational data – without ceding any integrative role to language production.

**Accounting for the correlational data.** As Carruthers correctly observes, the main correlational finding is that, of a range of variables – including ones for age, nonverbal IQ, verbal working memory and comprehension of spatial vocabulary – Hermer-Vazquez and her colleagues have succeeded in finding only one that correlates both positively and significantly with performance on the reorientation task, namely, performance on a left-right language *production* task (Hermer-Vazquez et al. 2001). To what extent does this finding support the restricted thesis? Here there appears to be a rather puzzling difference of opinion between Carruthers on the one hand and Hermer-Vazquez and her collaborators on the other. According to Carruthers: "Even by themselves, these data strongly suggest that it is language which enables older children and adults to integrate geometric with non-geometric information into a single thought or memory" (sect. 5.2). In contrast, according to Hermer-Vazquez et al. (2001), these "findings . . . are merely correlational, leaving open multiple possibilities for the relationship between developmental changes in spatial language production and advances in spatial memory task performance" (Hermer-Vazquez et al. 2001, p. 294). To mention merely one possibility, it may be that performance on the reorientation and production tasks *both* depend on a common developmental factor, such as the increasing accessibility of the concepts LEFT and RIGHT to central processes external to the geometry module. This and numerous other suggestions – some of which are considered in Hermer-Vazquez et al. (1999) – seem perfectly adequate to explain the correlational data without attributing any integrative function to language production. Why, then, is Carruthers so sanguine about the data?

Of course, the situation would be altogether different if Hermer-Vazquez had found that other variables – such as those for nonverbal IQ and language comprehension – are significantly and *non*positively correlated with reorientation performance. Under such circumstances, one's theory would need to account for the reason these variables did not correlate positively. And although I think that such an explanation could readily be provided without rejecting the traditional view, it would clearly be a harder task than merely explaining the production-reorientation correlation. In any case, this is decidedly *not* the situation. What, in fact, happened was that Hermer-Vazquez *failed* to find any other significant positive correlations; and this was simply because none of the other correlations – positive or otherwise – were significant. But – and this is the crucial point – under *these* circumstances, there is no need whatsoever to explain why the other variables do not positively correlate with reorientation performance. All one needs do is explain why the correlations are not significant; and there is an obvious explanation of this fact that in no way requires the endorsement of the restricted thesis: The sample was very small ($N = 24$).

To summarise (and at the risk of belabouring the point): Failing to find significant, positive correlations is *not* the same as finding a statistically significant absence of positive correlations. In particular, mere failures do not demand explanations of the absence, whereas significant findings do. It is perhaps, in part, for such reasons that Hemer-Vazquez and her collaborators seem inclined to view the dual-task data as providing stronger support for the restricted thesis.

**Accounting for the dual-task (shadowing) data.** The main dual-task finding is that verbal shadowing – but not rhythm shadowing – disrupts performance on the reorientation task. To what extent does *this* finding support the restricted thesis? Again – though for very different reasons – I'm inclined to think that it does not. To see why not, however, it's helpful to start by mentioning two assumptions about language production that are widely held among psycholinguists: (1) The language production system takes as input *communicative intentions,* and (2) *central* systems are responsible for combining – or *integrating* – representations to produce such intentions (Levelt 1989). If these assumptions are correct, then there is a natural four-step explanation of the dual-task data that cedes no integrative role to language production:

1. Given the above assumptions and given that verbal shadowing appears to be a *production* task, it's very plausible that successful verbal shadowing requires the activity of at least some of those central processes responsible for the construction of communicative intentions.
2. In contrast, it is very hard to see why *rhythm* shadowing should require the activation of such processes because there is no reason whatsoever to suppose that rhythm shadowing requires conceptual construction/integration of any sort.
3. Though little is known about the central processes responsible for conceptual integration, it is very plausible that – like *any* central process about which anything is known – they are subject to resource limitations. Moreover, it would be unsurprising if, because of such limitations, the simultaneous performance of two integrative tasks would tend to result in reduced performance on one or both.
4. Finally, suppose that the integration of geometric and nongeometric information in the reorientation task requires the activation of some of the *very same* central processes mentioned in (1).

Notice that the conjunction of (1) through (4) provide an explanation of the shadowing task data solely in terms of the activity of central processes. On the present story, verbal shadowing affects reorientation performance by loading some of the same central processes that are required to solve the reorientation task. In contrast, rhythm shadowing has no such effect because it simply does not require conceptual integration. Further, none of the four claims seem particularly tendentious. Rather, they are largely consequences of standard views about the nature of central systems and their role in language production and representational integration. Finally, it is worth noting that although compatible with the view that there are radically *amodular* central systems, nothing about the preceding proposal *requires* that we posit such mechanisms. So, for example, that explanation is entirely compatible with Sperber's (1996) suggestion that the kinds of conceptual integration involved in these tasks depend on what he calls a "metarepresentational module."

To summarise: The available data provide us with no reason whatsoever to endorse the restricted thesis that language production processes – as opposed to central processes – are responsible for the integration of geometric and nongeometric information in spatial reorientation tasks. To this extent, then, the best (putative) direct evidence for the fundamental thesis provides us with no reason to adopt it in favour of the competing traditional view of conceptual integration.

## Bilingual inner speech as the medium of cross-modular retrieval in autobiographical memory

Robert W. Schrauf

*Buehler Center on Aging, Northwestern University, Chicago, IL 60611-2611.*
**r-schrauf@northwestern.edu http://www.northwestern.edu/aging/**

**Abstract:** Carruthers's notion that natural language(s) might serve as the medium of non-domain-specific, propositionally based inferential thought is extended to the case of effortful retrieval of autobiographical memory among bilinguals. Specifically, the review suggests that the resources of bilingual inner speech might play a role in the cyclical activation of information from various informational domains during memory retrieval.

There is evidence that bilinguals recall their memories in the language(s) in which they were encoded (Marian & Neisser 2000;

Schrauf 2000). Particularly in the case of consecutive (immigrant) bilinguals, memories from the homeland seem to come back in the first language; memories from after immigration come back in either the first or second (Larsen et al. 2002; Schrauf & Rubin 1998; 2000). How are we to understand the relation between language (the natural languages of the bilingual) and thought (the process of autobiographical retrieval)?

On the one hand, if language has no constitutive role in thinking, but serves only for the outward narration of memory (Carruthers's communicative conception of language), then language(s) would be involved only in labeling features of an experience after it had been retrieved from memory where it is stored in some nonlinguistic form. Current models of bilingual memory in fact commit us to something like this position. These models describe how words in either of the bilingual's lexicons map onto underlying nonlinguistic conceptual and imagistic referents (Kroll & de Groot 1997). It is arguable, however, that autobiographical memories are very different from the atomistic nonlinguistic referents of lexical memory, but rather are complex, culturally and linguistically conditioned, narrative wholes at higher levels of mental organization (Schrauf & Rubin, in press).

On the other hand, if language is wholly constitutive of thinking (Carruthers's strong cognitive conception of language), then the bilingual's encoding and retrieval would be dramatically affected and memories encoded in one language might be only imperfectly retrieved in the other language (Schrauf 2000). Put bluntly, the immigrant would be amnesic in his or her second language for events that occurred in the first language in the homeland. Clearly, this is not the case, and between the former position and this, there must be some middle ground.

Carruthers's position provides a more promising foundation for bilingual autobiographical memory. The central process modularism underlying his theory fits the current understanding of autobiographical retrievals as complex integrations of multimodal sensory and conceptual information from various sites in the brain. And, at least in cases of conscious, effortful retrieval (e.g. "think of a time when you . . ."), natural language in inner speech may well function as the medium of non-domain-specific, cross-modular, propositional thought.

First, current conceptions of autobiographical memories as "on-line" reconstructions of events fit well with central process modularism. Autobiographical remembering is a temporally durative process of integrating different kinds of sensory and conceptual information from different brain sites (central and peripheral modules) into an episodic whole (Conway 2002; Rubin 1998). Information is processed at several sites: visual-object imagery (in occipital and parietal lobes), visual-spatial imagery (in the occipital lobes through inferior temporal cortex), auditory imagery (temporal cortex), emotion (limbic system), and conceptual information across the neocortex (Greenberg & Rubin, in press). Sensory modalities are to some extent informationally encapsulated and domain-specific (Moscovitch 1992). Both at encoding and retrieval, the medial temporal lobes are active in "binding" together the information (Schacter et al. 1998; Squire 1992); and at retrieval, when memory search is initiated in the frontal lobes (Wheeler et al. 1997), the same sensory and conceptual associations between sites are again activated. This domain-specific encoding of sensory and conceptual information, and coordination of information between domains, is consistent with the kind of central process modularism advocated by Carruthers.

Second, it might be that natural language plays a coordinating role in the inferential processes of effortful retrieval. Carruthers argues that in cases of inference, let us say, "problem-solving" (e.g., spatial cognition; see Hermer-Vazquez et al. 1999), central processing modules deploy the resources of the peripheral language module to coordinate information between domain-specific modules. In this exercise of propositional thought, natural language syntax provides a vehicle for the nested embedding necessary in organizing the information. Were it possible to conceptualize the integration of information in autobiographical retrieval in these same terms (as in propositionally based problem solving), then natural language might play a critical role in reconstructive retrieval as well. But that is precisely how effortful retrieval is described – as problem solving.

Effortful retrieval is a stepwise process in which information provided in the cue is used to construct search criteria, which trigger information in memory, which is evaluated against the criteria and reemployed to trigger further information in a cyclic fashion until the criteria are met and the final autobiographical memory comes together (Conway 1996). Behavioral evidence for this comes from protocol analysis in which individuals are asked to "think out loud" while they search for a memory matching the cue requirements (Williams & Hollan 1981; Williams & Santos-Williams 1980).

This cyclic search process through the structure of autobiographical memory is to some extent conscious and lends itself to inner speech (though much remains unconscious). Bilinguals are in fact aware that their inner speech can take place in either language: self-talk, calculating, thinking-in-words, praying, and so forth (Larsen et al. 2002; Schrauf & Rubin 1998); and although the "think-aloud" exercise is in overt speech, when bilinguals perform this exercise, they occasionally switch from one language to another for the strategic purpose of triggering (activating) new information (Schrauf, in press). Thus, reading the phenomenon of bilingual inner speech – the representation of propositional thought in the auditory imagery of two natural languages and the possibility of mental code switching – into effortful reconstructive retrieval offers new possibilities for testing the theory of natural language as a medium of interdomain thought and inference. For example, by choosing languages that are syntactically very different, we could investigate whether different structures of nesting and embedding information affect the amount or kinds of detail recalled from memories encoded in either language. In turn, accessing the natural-language inputs and outputs of the language module would require more sensitive methods than protocol analysis and thus more creative experimental design.

In sum, the notion that natural language might serve as a kind of *lingua franca* in the brain in service of propositionally based, problem-solving tasks sets the memory theorist a new way of conceiving the structure of bilingual cognition. In turn, probing the function of bilingual inner speech in reconstructive retrieval provides a way of testing and refining this version of the language-thought relation.

ACKNOWLEDGMENT
Research is supported by National Institute on Aging grant R01 AG16340.

## Talking to ourselves: The intelligibility of inner speech

Peter P. Slezak
*Program in Cognitive Science, School of History and Philosophy of Science, University of New South Wales, Sydney NSW 2052, Australia.*
**p.slezak@unsw.edu.au**
**http://www.arts.unsw.edu.au/sts/peter_slezak.html**

**Abstract:** The possible role of language in intermodular communication and non-domain-specific thinking is an empirical issue that is independent of the "vehicle" claim that natural language is "constitutive" of some thoughts. Despite noting objections to various forms of the thesis that we think in language, Carruthers entirely neglects a potentially fatal objection to his own preferred version of this "cognitive conception."

On Carruthers's own account, the empirical evidence he adduces only supports the essential involvement of language in certain kinds of thinking and not the quite different "constitutive" doctrine that natural language is the medium or representational vehicle of such thinking. Carruthers fails to consider an important and notorious objection to his own preferred version of the constitutive "cognitive conception."

As Carruthers acknowledges, the idea that we think "in" language in some sense is unfashionable today. However, despite the explanatory virtues of a rival Fodorian "mentalese," the idea that we think in natural language has a persistence and undoubted persuasiveness that needs to be addressed.

Even critics such as Jackendoff (1997), Pinker (1994), and Magee (1997) feel the need to address the idea that we think in language if only to refute it. However, these rebuttals are not new and have obviously failed to lay the doctrine to rest.

Ironically, Carruthers's own account suggests that this persuasiveness is because of a certain notoriously seductive illusion (Slezak 2002a; 2002b; in press). Carruthers's effort to revive and defend this cognitive conception of language has the virtue of clearly articulating the doctrine in its various guises, thereby exposing both the source of its attractiveness and its fatal flaws.

What does it mean to go beyond merely asserting the involvement of language in certain kinds of thought to claiming that natural language is constitutive as the medium or vehicle of our conscious propositional thoughts? The precise nature of this claim and the evidence for it are only referred to obliquely in Carruthers's talk of an "imaged natural language sentence." Indeed, from other writings (Carruthers 1996, p. 49) it is clear that Carruthers sees the usual kinds of scientific evidence for the constitutive claim as equivocal and, accordingly, he appeals to what he calls "a very different set of data, namely that delivered by ordinary introspection" (1996, p. 49).

Thus, for Carruthers, conscious thinking is constitutively in natural language in the sense that such thinking involves "imaging sentences of natural language." That is, sentences of natural language are recalled and briefly held in short-term memory (1996, pp. 228–29). Of course, this is the model of seeing with the mind's eye transposed to a different modality of imagination – hearing with the mind's ear, so to speak. Undeniably, introspection suggests that, just as we appear to be looking at pictures when we imagine visually, so we appear to be talking to ourselves when we think.

Thus Carruthers's thesis amounts only to the truism that when we are aware of "talking to ourselves" – the inner speech of folk psychology – we are simply imagining sentences of natural language, just as we might imagine anything else. However, such a claim does not amount to a theory of anything and, above all, has no implications for the nature of an underlying medium or substrate of thought. It is simply the avowal of introspection, just as one might report imagining anything else visual, tactile, or auditory.

U. T. Place (1956) famously observed that it is a "phenomenological fallacy" to suppose that when someone describes an experience of how things look, sound, smell, taste, or feel, then "he is describing the literal properties of objects and events on a peculiar sort of internal cinema or television screen" (Place 1956, p. 37).

Just as Place noted we are not likely to have anything green in our heads when we have a green afterimage, more recently, Edelman (1998) suggests that when we imagine a cat, we are unlikely to ascribe furriness to the mental representation. However, this very mistake seems remarkably compelling, though no more defensible, in the case of other attributes. Of course, in these remarks Place anticipated what Dennett (1991) has made well known as the fallacy of the Cartesian Theatre.

Significantly, Carruthers charges critics of the format or medium thesis with denying the stream of inner verbalization, or at least ascribing "some more-or-less epiphenomenal character" to it. It is revealing that this is the same strawman charge that has been leveled against critics of pictorial images (Kosslyn 1980, p. 30) – namely, that these critics deny the reality of imagery experience. Thus, it is no surprise, and no accident, that Carruthers invokes Kosslyn's account to support his own analogous thesis. However, it is remarkable that Carruthers makes no mention of the notorious controversy surrounding Kosslyn's theory. Of course, Kosslyn's pictorial account has been the focus of the intense "imagery debate" (Pylyshyn 1973; 1978; in press) and has been seen as the paradigm case of the Cartesian Theatre fallacy. Thus, with his appeal to pictorialism, Carruthers clarifies the nature of his own doctrine and at the same time brings into relief its fatal difficulties.

These parallels suggest that the issue of language and thought provides yet another case study of a pervasive mistake in reasoning about the mind found throughout cognitive science today and also throughout the history of the subject (Slezak 2002b).

Like pictures, sentences have their appeal as a medium for representation because we, as theorists, can understand them. But this is an illicit appeal to their intelligibility rather than their explanatory force. It is, of course, precisely the same illicit criterion employed by Searle to judge whether symbols can subserve intentionality. Their meaningfulness to Searle in the Chinese room is spuriously taken to be relevant to the theoretical question of meaning for the system. In effect, this is to require that the language of thought be English, just like Carruthers's thesis. However, the vehicles of meaning cannot themselves be meaningful in this sense on pain of the traditional regress of the homunculus pseudo-explanations.

Overlooked by critics and defenders alike is the fundamental question of the very intelligibility of the idea that we think *in* natural language. To raise this question is not merely to dispute an empirical claim but to challenge the very coherence of the doctrine. The point was made by Gilbert Ryle (1968) who challenged the very idea that "when thinking occurs, there occur internally or externally, things or symbols that the thinker thinks in" (1968, p. 12). This is not a behaviorism that would deny a Fodorian mentalese, above all, nor is it to deny that we do seem say things *sotto voce,* talking to one's self, as it were, in a silent monologue. Ryle attacks precisely Carruthers's doctrine, saying: "I want to deny that it even makes sense to ask, in the general case, what special sort or sorts of things we think in. The very collocation of "think" with "in so and so" seems to me factitious" (1968, p. 13).

## What language is "mentalese"?

Dan I. Slobin

*Department of Psychology, University of California, Berkeley, CA 94720-1650.* **slobin@socrates.berkeley.edu** **http://ihd.berkeley.edu**

**Abstract:** Carruthers's "mentalese expressions" take the form of English sentences, thus suggesting an isomorphism between thought and language that ignores linguistic diversity. Furthermore, complex syntax is not the only linguistic means of combining information from various domain-specific modules into domain-general expressions, nor is such syntax the preferred means of encoding basic experiences in all languages. The analysis seems to rest on an unacknowledged version of linguistic determinism.

I will comment on only one part of Carruthers's thoughtful and stimulating paper. While I quite agree that language should not be treated simply as "a channel, or *conduit,* for transferring thoughts into and out of the mind," I think it is just as problematic to treat language as a faculty that takes "thoughts . . . *for direct translation into natural language expression"* (emphasis added). This move requires a common format for thoughts and linguistic expressions – in effect, turning "thought" into a linguistic system of representation. Carruthers's procedure is to use small capitals for "mentalese expressions," while ignoring the fact that these expressions are literal translations of *English* sentences. One cannot say, in every language, THE TOY IS IN THE CORNER WITH A LONG WALL ON THE LEFT AND A SHORT WALL ON THE RIGHT. That is, this is not a pure mentalese expression that can be "directly translated" into a natural language. For example, if you speak a language that does not use left and right to locate objects but rather compass points or relation to landmarks (Levinson 1996; 1997; 2002), you would never have the mentalese expression just given. Your language would give you translations of mentalese expressions using phrases such as "the long wall to the North" or "uphill" or "seaward" or whatever. The research of Levinson and his colleagues shows quite clearly that users of different languages store environmental information with regard to the distinctions called for in speaking that language. So, for example, Australian aborigines can accurately point to distant locations and to compass points, whereas Europeans tend to be hopeless on similar tasks. Or they can quite readily warn you of a mosquito approaching your north arm when you may have no clue which arm is "north." Mentalese, for the Australian, would have to be in the form of, for example, THE TOY IS IN THE CORNER WITH A LONG WEST WALL AND A SHORT EAST WALL. Further, of course, relations such as IN, locations such as CORNER, and so forth, are specific to particular languages (Bowerman & Choi 2001; Choi & Bowerman 1991) and can't be part of mentalese. I will not further summarize the growing body of evidence for the ways in which particular languages "reach back" into the conceptual system to bias speakers toward language-specific ways of encoding basic experiences. Carruthers reviews this research in section 2.2 and seems to accept some version of "language as sculpting cognition." (Also see discussions of "thinking for speaking" in Slobin 1996; 2003.)

Where Carruthers seems to go astray is in his undefined uses of the terms *sentence* and *language.* For him, "the thoughts generated by central modules are used to produce a natural language *sentence* with the *same content.*" What is "the same content" for a thought and a sentence? If one thinks of language as only *one* language – say, English – the problem looks tractable. But even here it is illusory, because there is no way to line up the format of thoughts and sentences without pushing the form of sentences back into the form of thoughts. (In addition, any situation is open to an indefinite number of different linguistic encodings in any individual language.) And when one considers even two different languages, the task becomes impossible. There is no neutral "thought" (in capital letters) that translates into "wall on the left" as well as "west wall" and "uphill wall" and "seaward wall." That is, the user of mentalese must already know which system of orientation is called for in the local language. In brief, Carruthers has replaced a notion of language as a channel of communication with the notion of language as a notation system for thoughts. Without anchoring thought and language, however, this seems to be a hopeless enterprise.

At the next stage in his model, language combines the sentences translated from several types of thoughts into multiclausal sentences, such as, "The food is in the corner with the long wall on the left by the blue wall." That is, the role of language in human cognition is to combine "distinct domain-specific sentences . . . into a single domain-general one," thereby allowing for conscious, non-domain-specific thinking. Is Carruthers claiming that such sentences are necessary, in consciousness, to accurately remember where the food is? Introspection, along with observation of child language, hardly supports such a proposal. These are not the everyday forms of inner speech. Further, Carruthers's discussion of aphasia shows that such thinking can be conducted without the ability to produce sentences. Therefore I am not convinced that conscious problem solving requires the use of certain types of complex sentences in inner speech.

So, if these sentences are not conscious productions of inner speech, what are they? Carruthers requires these mental processes to be conscious, yet I don't find them in my consciousness. Is the next step to propose that the format of my *pre*conscious or *un*conscious thought is complex English sentences, provided by a language faculty that is equipped with translations of domain-specific propositions in mentalese? But consider languages that analyze situations and events in quite different terms from English. That is, the same information can be presented in *separate* clauses – in fact, this is *required* in many languages: For example, "There's some food. It's over there by that wall – the long one, by your left hand – the blue one. It's where the long wall and the other wall come together." And so forth. See, for example, Pawley's (1993) analysis of event descriptions in Kalam, a language of New Guinea that requires analysis of events into a chain of separate segments, each encoded by a separate verb with no embedding, or Bohnemeyer's (2003) discussion of Yukatek Mayan, in which one cannot say, "He went from his house to town," but is required to say "He was at his house and then he left and then he arrived at town." In brief, complex thoughts do not require complex syntax (although they may well require complex discourse structures).

Carruthers seems to propose that the format of mentalese is isomorphic with the format of English. Or, if he accepts cross-linguistic diversity, then the format of mentalese matches the format of the particular language in which it is to be translated. Either way, we seem to be faced with an untenable version of linguistic determinism.

## Conjoining information from different modules: A comparative perspective

Giorgio Vallortigara[a] and Valeria Anna Sovrano[b]

[a]*Department of Psychology and B.R.A.I.N. Centre for Neuroscience, University of Trieste, Trieste, Italy;* [b]*Department of General Psychology, University of Padua, Padua, Italy.* **vallorti@univ.trieste.it** **sovrano@mail.psy.unipd.it**

**Abstract:** The hypothesis that nonhuman species, lacking verbal language, do not really integrate information from different modules, but use instead information sequentially, appears difficult to put under empirical scrutiny. Evidence is discussed showing that in nonhuman species storing of geometric information occurs spontaneously even when landmark information suffices for spatial reorientation, suggesting simultaneous encoding, if not use, of information from different modules.

We would like to discuss some comparative data that are relevant to Carruthers's hypothesis that language serves as the vehicle of intermodular integration. Carruthers mentioned the challenge provided by research with young chickens (Vallortigara et al. 1990) and monkeys (Gouteux et al. 2001) suggesting that nonhuman (and thus nonlinguistic) species can apparently integrate geometric and nongeometric (landmark) information when disoriented. Actually, there are two other species that have been tested in the reorientation task originally developed for rats (Cheng 1986), namely, pigeons (Kelly et al. 1998) and, very recently, fish (Sovrano et al. 2002); both these species have been proved to be able to integrate geometric and landmark information to reorient themselves. The case of fish is particularly intriguing, for they have been tested in the same blue-wall task in which children have been shown to rely exclusively on geometric information (Hermer & Spelke 1994; 1996).

Carruthers provides a very ingenious way to face the potential challenge offered by these comparative data on the hypothesis that integration of information from different modules would not be possible without language. His idea is that even though other species are able to solve the spatial disorientation problem, this does not prove that they are able to integrate geometric with landmark information into a single belief or thought, for it could be that they are making use of the information sequentially. According to this view, the difference between species such as chicks, pigeons, rhesus monkeys, and fish, on the one hand, and rats and (prelinguistic) children, on the other hand, would simply be that the former use nongeometric information first before using geometry, whereas the latter use geometry exclusively.

We are sympathetic with Carruthers's ideas, but we have difficulty in understanding how they can be tested empirically. From the behavioral point of view, we simply cannot distinguish between the performance of the linguistic and nonlinguistic creatures that solve the spatial reorientation task. Thus, saying that nonhuman animals are not "really" able to integrate information from the geometric and the landmark modules is just a *petitio principii,* until we are able to specify for what aspects the spatial reorientation behavior of chickens and fish is different from that of human adults. Evidence that preventing verbal encoding in human adults impairs the capacity to integrate geometric and landmark information (Hermer-Vazquez et al. 1999) does not prove anything with respect to the abilities of nonhuman animals. It could be that the novel-acquired, human-specific, language abilities interfere with the spatial modules we share with other animals, but that integration between spatial modules can occur anyway even in the absence of language. If language is so crucial to the integration of information from different modules, then there should be quite specific difficulties that nonhuman animals should exhibit in the spatial reorientation tasks. The problem is that, at present, we are unable to unravel them. What sort of test could disprove the idea that animals can *really* integrate geometric and nongeometric information in a single thought?

Let us consider some recent results obtained in our laboratory to explain the nature of the difficulty (Sovrano et al. 2003). We trained fish with the usual rectangular arena (a rectangular tank in this case) with four distinctive panels located at the corners. In this case no integration of geometric and nongeometric information is needed to solve the reorientation task: Each corner can be identified without any ambiguity on the basis of solely the featural information provided by the panels. Surprisingly, we found that the fish managed the task. According to Carruthers's hypothesis, they should possess (like chickens and monkeys) an innate predisposition to seek landmark information first, only using geometric information to navigate in relation to a known landmark. However, given that in this case geometric information was not necessary for reorientation, we should probably have expected that no encoding of such information would occur. Surprisingly, when tested after removal of all the panels, fish did not go back to random choice, but searched systematically in the two locations specified by purely geometric information (similar results have been obtained with chicks; see Vallortigara et al. 1990). This clearly provides evidence that fish encode geometric information even when not required, that is, even when featural information suffices to solve the task. But does it provide evidence that fish had encoded geometric and nongeometric information in a single belief? Probably not, for it can be claimed that encoding of geometric and nongeometric information by two separate modules occurs anyway, even when this is not required by the task, but that information from these modules can then be *used* only sequentially. Thus, the data cast doubts over the adaptationistic account provided by Carruthers, because it seems that there is a primacy of geometric information even in those species (like fish and chickens) that solve the spatial reorientation task. Nonetheless, the data do not disrupt Carruthers's central tenet, for the crucial hypothesis is that information is used sequentially, but the order in which information is used can well vary in different species (though, admittedly, it remains mysterious that rats use only *one* type of information). Carruthers's hypothesis thus does not seem to be easily viable to falsification. It seems that whatever behavioral performance we could document in animals, it would always be possible to explain it as a result of a sequential, rather than a simultaneous, use of information from different modules.

Cognitive penetrability of a module can be assessed on the basis of behavioral performance. The behavior of at least some nonlinguistic animal species and of human adults shows that nongeometric information is taken into account in the spatial reorientation task. Hence we should conclude that language is not needed to integrate information from the geometric and landmark module. Alternatively, we should prove that, even if performance in this task is identical, animals fail in other aspects of the spatial reorientation process. In other words, we should prove that there are aspects of spatial behavior of humans that are not accessible or possible to nonhumans because of a supposed crucial difference, namely, a lack of possession of a verbal language allowing intermodular integration. Otherwise, the supposed difference would not make any difference.

## Language, cognition, and the nature of modularity: Evidence from aphasia

Rosemary Varley[a] and Michael Siegal[b]

[a]*Department of Human Communication Sciences, University of Sheffield, Sheffield, UK S10 2TA;* [b]*Department of Psychology, Western Bank, University of Sheffield, Sheffield, UK S10 2TP.* **R.A.Varley@sheffield.ac.uk M.Siegal@sheffield.ac.uk**

**Abstract:** We examine Carruthers's proposal that sentences in logical form serve to create flexibility within central system modularity, enabling the combination of information from different modalities. We discuss evidence from aphasia and the neurobiology of input-output systems. This work suggests that there exists considerable capacity for interdomain cognitive processing without language mediation. Other challenges for a logical form account are noted.

Peter Carruthers provides an elegant tour across the landscape of competing claims on the role of language in human cognition. This in itself is a complex undertaking, in view of the still limited evidence. For example, the evidence from aphasia only provides a window on the role of language in the mature cognitive system and says nothing as to how language might sculpt a distinctively human cognitive architecture. Hence a whole class of "diachronic" claims regarding the effects of language on cognition lie outside of the scope of the evidence from aphasia. Moreover, the evidence from aphasia studies of either associations or dissociations between language impairment and other cognitive functions should be interpreted with caution. Evidence of associations – for example, between aphasia and an inability to generate alternative strategies on a problem-solving task (Varley 2002; Varley et al. 2001) – falls far short of what is needed to draw causal inferences. Associations may result from naturally occurring brain lesions crossing the boundaries of the neural substrates of processing systems; the association is then an anatomical coincidence rather than a causal relationship in which language gives rise to cognition. Equally, dissociations between language impairment in aphasia and the sparing of cognitive performance in other domains such as space and number require careful consideration. In many studies, either the nature and extent of linguistic impairment are not clearly specified, or the abilities that are specified are such that the aphasic individuals have considerable residual language capacity still available to them that could support cognition. This capacity would be sufficient to generate simple "proto-language" – a language that is characterized by combinatorial properties but does not have the full features of a mature grammatical system (Bickerton 1981).

The proposal of massive central system modularity requires a mechanism that allows flexibility between domains of cognition. Although the notion of modularity is subject to different interpretations (e.g., Coltheart 1999; Fodor 1983), it is widely acknowledged that this must involve core characteristics such as domain specificity, informational encapsulation, and a proprietary operating code. However, flexibility and fluidity characterizes human cognition, particularly in higher-level reasoning and problem solving; hence the requirement for a cognitive *lingua franca* to combine the outputs of the modular subsystems. For Carruthers, the central code that permits the combination of outputs is logical form – the abstract, deep structural representations that underlie natural language sentences.

There are a number of possible reasons to reject the claim that natural language in the form of logical form sentences is the mechanism for intermodular combination. For example, the nature of the brain architecture of input-output systems permits intermodal mixing at an early stage in information processing. Sensory-perceptual systems consist of batteries of micromodules that respond only to very specific inputs (e.g., the tonatopic organization of the primary auditory cortex). However, zones of association cortex, described as polymodal areas, surround primary sensory cortices. Association cortex receives inputs from a number of micromodules within a sensory domain such as vision or from different sensory domains. Tasks such as those of Hermer-Vazquez et al. (1999) that require the combination of visual object and geometric information are central to Carruthers's thesis on the relation between language and cognition. However, responses in such tasks could be achieved by a lower-level nonlinguistic perceptual mechanism. Evidence from language training studies that result in accurate spatial cognition could reflect a co-opted strategic resource that is sufficient to assist performance on a range of problem-solving tasks, including theory of mind (Siegal & Varley 2002). There is debate about the extent of association cortex, with the suggestion that as perception is gradually fractionated into further micromodules, more of what is believed to be polymodal cortex may be reclassified as primary, sensory cortex (Kaas 1999). However, it is well established that the architecture of sensory-perceptual cortices permits the integration of information from different domains at an early stage in cognitive processing.

Another neurobiological example of interdomain information combination involves mirror neuron systems. These are cells in motor regions that fire when an action is observed, thus providing a mechanism for input-output linkage again at a low level of information processing (Rizzolatti et al. 2001). Both polymodal association cortex and mirror neuron systems show that the brain is equipped with mechanisms to permit combination of interdomain information at an early stage in information processing. These neural systems are intrinsic to input-output modules and are remote both anatomically and functionally from central cognitive systems.

There are other difficulties at both the theoretical and empirical levels for an account based on central modularity. For example, Carruthers's account suggests that the language faculty generates logical form (LF) sentences in response to inputs from a range of central modules. This is a problem because now a fundamental criterion for modularity appears to be violated in that the language production system is able to take as input a multiplicity of codes from a range domain-specific modules. The language faculty, or at least its output component, then becomes an omnipotent domain-general system, and Carruthers's attempt to develop the thesis of central system modularity results in a breakout of domain generality elsewhere in cognition.

We conclude by noting that an LF account poses a considerable challenge for experimental research. Imagine a set of results from a Hermer-Vazquez et al. type of task in which one subgroup of people with aphasia shows the ability to combine landmark and geometric information whereas others do not. In our laboratory, we have carried out a pilot study involving persons with severe aphasia in which the Hermer-Vazquez et al. room procedure is scaled down to a table-top box. In this task, rather than disorientating the patient, the box is spun. Our subjects were capable of utilizing both landmark and geometric cues in locating a hidden object, pointing to the sparing of their spatial cognition. They could find objects under conditions of no concurrent activity as well as in a simple lexical verbal shadowing condition where residual lexical capacity was engaged by the concurrent activity. As the aphasic participants had a profound impairment of their public language, a counter to these results could be that, at a deep and underlying level, the LF representation was intact. It is difficult to determine how the integrity or otherwise of LF representations could be established, given that evaluation of language processing at both implicit and explicit levels requires the capacity to process surface form. Similarities in aphasic impairments across input and output modalities may be seen to indicate a loss of central language competence. But this inference is an imprecise and indirect metric of the integrity of logical form. An alternative account might postulate parallel input-output impairments – resulting in disturbance of public language that would mimic a central competence deficit. Thus a crucial element in the evaluation of Carruthers's model must be whether it generates testable hypotheses.

## The role of working memory in skilled and conceptual thought

Thomas Wynn[a] and Fred Coolidge[b]

[a]*Department of Anthropology, University of Colorado, Colorado Springs, CO 80933-7150;* [b]*Department of Psychology, University of Colorado, Colorado Springs, CO, 80933-7150.* **twynn@uccs.edu** **foolidg@uccs.edu**

**Abstract:** Models of working memory challenge some aspects of Carruthers's account but enhance others. Although the nature of the phonological store and central executive appear fully congruent with Carruthers's proposal, current models of the visuo-spatial sketchpad provide a better account of skilled action. However, Carruthers's model may provide a way around the homunculus problem that has plagued models of working memory.

Carruthers's review of the various hypotheses of the cognitive functions of language would have been aided invaluably by reference to Baddeley's provocative and influential model of working memory (Baddeley 1986; 2001; Baddeley & Logie 1999) and vice versa.

For example, most current models of working memory (Miyake & Shah 1999) suggest that some coordination of separate cognitive domains takes place independent of language. The phonological store of working memory is well documented experimentally and is clearly the neural system responsible for holding words in mind during the construction and comprehension of sentences. Similarly, the visuo-spatial sketchpad (VSSP) holds visual and spatial information within immediate attention. Work by Logie (1995) now indicates that spatial information (e.g., location and direction) and visual information (e.g., color and shape) are actually separate subsystems of the VSSP. This interpretation is congruent with Kosslyn's (1994) description of visual imagery, in which shape and spatial information is seen to be coordinated by frontal lobe functions. So here we have evidence for interdomain coordination in the absence of language. This example might well be considered irrelevant (one could argue that the total output of the VSSP should be considered a single content domain) if it were not that it accounts for one of the central pieces of evidence put forward by Carruthers – the coordination of geometric properties (spatial) with color (visual). Carruthers suggests that children only solve the hidden object problem when they are able to use "spatial vocabulary conjoined with object properties." We suggest that this particular ability is more easily understood as being a classic VSSP task, and that children succeed at the problem when working memory has developed sufficient capacity. Results of dual interference tasks (e.g., Hermer-Vazquez et al. 1999) could also be accounted for by interference with the central executive of working memory rather than language per se, as argued by Carruthers.

Work in cognitive anthropology reinforces the idea that the VSSP is important to thinking but is largely nonverbal. Keller (Dougherty & Keller 1982; Keller & Keller 1996), for example, describes how individual steps in the blacksmithing process come together in "constellations" brought to mind and tied to specific intermediate plans of action. These constellations include not only actual tools but also visual images of shape and color (crucial to temperature judgments), aural images of the sound of forging (a measure of the malleability of iron), and kinesthetic images of muscle tensions, angle of grip, and so on. Little or none of this is represented verbally, nor indeed is it easy to elicit post hoc accounts from artisans, who "grope for words" to describe the process. One could of course argue that this is not conceptual thought, a solution hinted at by Carruthers when he grants "practical" reasoning to all mammals. Alternatively, one could argue that the inner speech is somehow outside the attention window of the smith, but it seems simpler to conclude that there are sophisticated human actions that are coordinated in a nonverbal way via the VSSP and central executive of working memory.

From this perspective, some non-domain-specific thinking appeared prior to the emergence of modern language. Archaeological evidence for coordination of visual and spatial information goes back at least 300,000 years (probably more) in the guise of handaxes ("handaxe" and "handaxes" are single words) with three-dimensional symmetry (Wynn 2000, in press). Manufacture of these artifacts required coordination of spatial thinking – allocentric perception, size constancy (Silverman et al. 2000) – and shape (symmetry), which is done via the VSSP of working memory.

We do not think, however, that the preceding arguments challenge Carruthers's overall thesis vis-à-vis the role of inner speech; they merely question its scope (all non-domain-specific thinking is based on inner speech) and one piece of its evidentiary base (coordination of visual and spatial knowledge). Otherwise, Carruthers's thesis meshes well with current understandings of the phonological store and central executive of working memory.

Phonological storage for Baddeley is a short-term buffer where sentences are prepared, rehearsed ("inner speech"), and expressed (outwardly articulated). Baddeley has recently proposed an episodic buffer (or non-domain-specific central buffer in Carruthers's language) that may be identical to the role Carruthers has proposed for inner speech. According to Baddeley, the episodic buffer briefly stores and may chunk output from other systems using different codes. Interestingly, Baddeley (2001) claims that the buffer may create an environmental model that can be manipulated "to solve problems and plan future behavior." Carruthers asks, "What are the consumers for the LF [Chomsky's logical form] sentences generated from the outputs of the central modules?" It appears that at least part of Carruthers's answer may lie in Baddeley's concept of an episodic buffer. But the parallels in Carruthers's and Baddeley's theories become even more compelling.

Carruthers proposes that in addition to a domain-general factual memory system (long-term, declarative memory in Baddeley's model), there must be an innate, domain-general faculty of abductive inference. He labels it a "consumer system," but it seems very like Baddeley's central executive. Baddeley views the central executive as largely attentional in nature; that is, it directs the cognitive functions of the mind (including language) to relevant tasks while inhibiting attention from irrelevant contexts. Carruthers proposes that "central cognition" uses language (an innately structured input and output module) when engaging "in certain kinds of reasoning and problem solving." Baddeley wrestles with the problem of a homunculus, that is, who or what actually makes the decisions on these "inner propositions." Carruthers hints at an answer: It is innate, universal, and appears developmentally early (by age 4). He cites the work of Frankish (1998a; 1998b), and Carruthers purports that the "distinctive causal role of inner speech is partly a function of our *decisions* to accept, reject, or act on the propositions which our imaged sentences express. . . . Then provided that I can remember my commitments and execute them, it will be *just as if* I believed the proposition in question" (sect. 4, para. 4). Carruthers presents Frankish's proposal that a level of our mentality is constituted by our higher-order decisions and commitments to accept or reject propositions, and also that language constitutes the thoughts and beliefs entertained at that level. Again, Baddeley's concept of a central executive is germane to these suppositions; although here is an example where Carruthers's speculations about the nature of this higher-order, innate abductive mode of thinking would have furthered Baddeley's speculations about the nature of the central executive.

## Language and conceptual development: Words as essence placeholders

Fei Xu

*Department of Psychology, Northeastern University, Boston, MA 02115.*
**f.xu@neu.edu**

**Abstract:** Perhaps in addition to language being a potential medium of domain-general thought, as suggested by Carruthers, language may also play another role in conceptual development: Words are "essence placeholders." Evidence is presented from studies on categorization, object individuation, and inductive inference in infancy. The assumption that words are essence placeholders may be a mechanism by which infants acquire kind concepts.

Carruthers argues for a particular version of the cognitive function of language in his target article, where language serves to conjoin modular representations. This is an intriguing thesis. However, the empirical evidence seems to be fairly weak, for example, children placed in a larger room succeeded on the reorientation task without the aid of language. I suggest a different role for language: Even at the very beginning of language development, words may serve as "essence placeholders" for human infants. The empirical evidence comes from studies in cognitive development, in particular, studies on categorization, object individuation, and inductive inference.

In a series of categorization studies, Balaban and Waxman (1997) showed that labeling can facilitate categorization in infants as young as nine months. In their studies, the infants were familiarized with a set of pictures of a given category, say, rabbits. One group of infants heard a word when shown the picture, for example, "a rabbit," on some trials. For a second group of infants, a tone accompanied the presentation of the picture on some trials. The results showed that although both words and tones heightened the infants' attention to the objects, it was only in the label condition that the infants categorized the objects. That is, they preferentially looked at an exemplar from a new category (e.g., a pig) compared to an exemplar from the old category (e.g., a new rabbit). These results suggest that in the presence of a label, infants group together the exemplars into a single category more readily than in the absence of a label.

In a series of experiments with 16- to 21-month-old infants, Welder and Graham (2001) found that labeling guides infants' inductive inference, and it sometimes overrides perceptual similarity. In their studies, the experimenter modeled a nonobvious property of an unfamiliar target object, for example, an object making a rattling sound when shaken, and the infant was given other test objects that were perceptually similar or dissimilar to the target object. In addition, one group of infants heard a label applied to both the target object and the test objects, whereas a second group did not hear a label. The question of interest was whether infants would predict that the test objects would also possess the nonobvious property by reproducing the action; that is, they should shake the object if they expected it to rattle. The results showed that labeling had a powerful effect on inductive inference, often overriding perceptual similarity.

A third line of experiments suggests that in addition to facilitating categorization and inductive inference, labeling may also play an important role in object individuation. Xu (2002) presented nine-month-old infant's a task in which an object, say, a toy duck, emerged from behind an occluder and returned behind it, followed by another object, say, a ball, emerging from behind the same occluder and returned behind it. The question was whether the infants can use the property or object kind information to conclude that there were two distinct objects behind the screen. The results showed that when the two objects were labeled by distinct words, that is, "a duck" and "a ball," the infants succeeded on the task, whereas when the two objects were labeled by the same word, that is, "a toy," the infants failed. The labeling effects held up even with unfamiliar objects and nonsense words, for example, "a blicket" and "a tupa." Subsequent experiments showed that this effect might be language specific because other auditory stimuli, for example, distinct sounds or emotional expressions (e.g., *ah* versus *eeu*), did not facilitate object individuation.

These results suggest that even early in language development, labeling may exert influence on several aspects of cognition. Infants may expect that words for objects map onto distinct kinds in their environment; words are "essence placeholders." Given this expectation, if a set of objects are all labeled as "a rabbit," it provides evidence to the infant that they belong to the same kind. Similarly, the very fact that one object is called "a duck" and one object seen on a different occasion is called "a ball" is sufficient evidence for infants to posit two distinct kinds or essences. According to psychological essentialism (Gelman 2003; Medin & Ortony 1989), essences determine the surface features and properties of objects. Three consequences follow. First, if two objects share a label, they should belong to the same kind and their featural similarities should be analyzed, as in the categorization task. Second, if two kinds of objects are indeed behind an occluder because two distinct labels are heard, there must be two distinct tokens of objects, as in the object individuation task. Third, if two objects share a label, they should have the same nonobvious properties, as in the inductive inference task. The assumption that words are essence placeholders may be a mechanism by which infants first establish what *kinds of things* are in their environment.

# Author's Response

## Modularity, language, and the flexibility of thought

Peter Carruthers

*Department of Philosophy, University of Maryland, College Park, MD 20742.*
**pcarruth@umd.edu**
**www.philosophy.umd.edu/people/faculty/pcarruthers/**

**Abstract:** The present response elucidates, elaborates, and defends the main thesis advanced in the target article: namely, that natural-language sentences play a constitutive role in some human thought processes, and that they are responsible for some of the distinctive flexibility of human thinking, serving to integrate the outputs of a variety of conceptual modules. Section R1 clarifies and elaborates this main thesis, responding to a number of objections and misunderstandings. Section R2 considers three contrasting accounts of the mechanism of intermodular integration. Section R3 discusses objections to some of the empirical data supporting my main thesis. Section R4 criticizes some competing theories of the role of language in cognition. And section R5 considers some proposed supplementary cognitive roles that language might play.

### R1. Clarifying and elaborating the main thesis

The main thesis of the target article is that natural language sentences, constructed by the language module, serve to integrate and combine the outputs of a variety of central-process conceptual modules, hence enabling some of the flexibility which is distinctive of human cognition. In this section, this thesis will be both clarified and elaborated on. Only by casting the motivation behind my main thesis in a new light, in fact, and by adding some later developments, can I put myself in position to reply to the detailed criticisms of my commentators.

#### *R1.1. On mental modularity*

A number of commentators challenge the assumption of mental modularity that forms the backdrop to my proposal (**Bickerton**, **Dale & Spivey**, **Hurford & Dessalles**, **MacWhinney**), and **Hermer-Vazquez** complains of the difficulty of defining what modularity *is.* Otherwise, most are neutral or express some degree of support. Only two, however (**Atran**, **Bryson**), identify the primary reason for believing in mental modularity: computational tractability. Let me elaborate on the positive before replying to the negative. Part of the point of the discussion will be to clarify exactly how I think the modularity thesis should be construed.

The most important argument in support of mental modularity derives from Fodor (2000). It goes like this:

1. The mind is computationally realized.
2. Amodular, or holistic, processes are computationally intractable.

Conclusion: So the mind must consist wholly or largely of modular systems.[1]

Premise 1 is the guiding assumption that lies behind all work in computational psychology, hence gaining a degree of inductive support from the successes of the computationalist research program. And many of us believe that this form of psychology represents easily our best hope (perhaps our *only* hope) for understanding how mental processes can be realized in physical ones (Rey 1997).

Just about the only people who reject Premise 1 are those who endorse an extreme form of distributed connectionism, believing that the brain (or significant portions of it, dedicated to central processes) forms a vast collection of connectionist networks, in which there are no local representations (**O'Brien & Opie**). The successes of the distributed connectionist program have been limited, however, mostly being confined to various forms of pattern recognition; and there are principled reasons for thinking that such models cannot explain the kinds of structured thinking and one-shot learning of which humans and other animals are manifestly capable (Fodor & McLaughlin 1990; Fodor & Pylyshyn 1988; Gallistel 2000; Horgan & Tienson 1996; Marcus 2001). Indeed, even the alleged neurological plausibility of connectionist models is now pretty thoroughly undermined, as more and more discoveries are made concerning localist representation in the brain (e.g., Rizzolatti et al. 2001).

Premise 2 is almost universally accepted, too, and has been since the early days of computational modeling of vision. You only have to *begin* thinking in engineering terms about how to realize cognitive processes in computational ones to see that the tasks will need to be broken down into separate problems that can be processed separately (and preferably in parallel). And indeed this is, for example, exactly what we find in the organization of the visual system.

What this second premise does, then, is to impose on proposed modular systems quite a tight *encapsulation* constraint. Any processor that had to access the full set of the agent's background beliefs (or even a significant subset thereof) would be faced with an unmanageable combinatorial explosion. We should therefore expect the mind to consist of a set of processing systems that are not only modular in the sense of being distinct isolable components but that operate in isolation from most of the information that is available elsewhere in the mind.

I should emphasize that this point concerns not the range of *inputs* that are available to modular systems, but rather the *processing data-bases* that are accessed by those systems in executing their algorithms (Carruthers 2003; Sperber 2002). Failure to appreciate this point leads **Varley & Siegal** to complain that language can't be a module, because it receives inputs from many other modules. Although we might *expect* some modules to receive only a restricted range of inputs, this is by no means a requirement of modularity.[2] And indeed, the practical reasoning module envisaged in the target article is hypothesized to receive inputs from all belief-forming and desire-forming modules (see also sect. R1.3). This is quite consistent with its modular status, provided that the information that can be accessed when processing any given set of inputs is strictly limited.

I should also emphasize that modularity is perfectly consistent with an acknowledgment of the importance of *attention* in organizing the direction of modular resources and also with a variety of kinds of top-down processing. (Both are found in the operations of the human visual system; see Kosslyn 1994.) Some kinds of attention can be thought of as directing the inputs to a module. And top-down processing can all be internal to a wider modular system. What is important, from a modularist perspective, is just that the processing within the system shouldn't require habitual access to significant quantities of information held elsewhere in the mind.

Modularism is now routinely assumed by just about everyone working in artificial intelligence, in fact (Bryson 2000; McDermott 2001). So anyone wishing to *deny* the thesis of massive modularity is forced to assume a heavy burden. It must be claimed either that minds *aren't* computationally realized, or that we haven't the faintest idea how they can be. And either way, it becomes quite mysterious how minds can exist in a physical universe. (This isn't to say that modularism doesn't have its own problems, of course. The main ones will be discussed briefly in section R1.4. And one of the major purposes of my target article was to make a start on addressing them.)

Any computational psychologist should predict, then, that the brain will contain a whole host of modular processing systems, connected with one another in a variety of complex ways, many of which can be decomposed into networks of smaller modules. (See any of the familiar and highly complex wiring diagrams of the visual cortex to get the flavor of the kind of thing I have in mind.) And nothing in this account need prevent one submodule (e.g., for processing faces) from forming a common part of a number of distinct higher-level modules either (e.g., not only the visual system but also an emotional system).

So far as I am aware, such a picture is consistent with (and to some degree supported by) everything we currently know about the organization of the brain. Certainly it is no problem for this sort of modularist account that brain-imaging studies show dispersed networks of neurons engaged in any given task, to which **Bickerton** draws our attention. First, most *tasks* are likely to involve the activity of a variety of modules, both input and output; and second, many modules may well be composed out of submodules that are located in different brain regions. (See, e.g., Baron-Cohen & Ring 1994 for a tentative brain map of the subsystems involved in our theory of mind module, or ToMM.)

Nor do the elegant eye-tracking data described by **Dale & Spivey** count against this sort of modularist account.[3] That language comprehension should be subserved by a complex system involving a language module receiving inputs from both speech and vision, and an attentional system for coordinating the two sorts of input (as well as a number of other parts besides, probably) is exactly what a modularist of my stripe might predict (see also Chomsky 2000; Pietroski 2003).

Is a modularist position implausible on evolutionary grounds in the way that both **Bickerton** and **Hurford & Dessalles** allege? In fact, no. What we know from biology generally is that evolution is apt to operate by constructing specific systems in response to specific adaptive pressures (Tooby & Cosmides 1992). And often this will happen by co-opting and reconfiguring resources that were available as antecedents, in such a way that multifunctionality becomes rife (**Hermer-Vazquez**). What *would* really be biologically implausible would be the evolution of a big equipotent brain (given the high costs of sustaining large brain size) or the evolution of an unstructured central processing arena, of the sort many antimodularists are committed to (Tooby & Cosmides 1992).

### *R1.2. Integration of module outputs without language*

A great many commentators object to my thesis that language is the medium of intermodular integration by pointing out cases where the outputs of different modular systems are integrated *without* language (**Bryson**, **Chater**, **Dale & Spivey**, **Hampton**, **Hermer-Vazquez**, **Hurford & Dessalles**, **Nelson**, **Robbins**, **Slobin**, **Varley & Siegal**, **Wynn & Coolidge**). For example, the shape and space modules in the visual system are integrated without language (**Wynn & Coolidge**), and the mirror neurons discovered by Rizzolatti et al. (2001) integrate both visual and proprioceptive information (**Bryson**, **Varley & Siegal**). But as will be shown by the account of mental modularity sketched in section R1.1, I, too, believe that modular systems will frequently receive inputs from two or more distinct modules, and hence that integration of modular outputs will frequently occur in the absence of language. Nothing that I said in the target article was intended to deny this. Rather, the thesis that I advanced there assumed a certain sort of overall mental architecture and was intended to be restricted to just one level of that architecture.

What I envisage is a layer of perceptual-input modules (including language) that generate outputs to be taken as input by a layer of conceptual belief-forming and desire-forming modules. The input modules may bear complex relations to one another (e.g., both hearing and vision feed into language, as noted in sect. R1.1) and may themselves decompose into a variety of submodules. And the conceptual modules, similarly, will bear complex relations to one another, as well as being further decomposable.

The conceptual modules in turn feed their outputs to a limited-channel practical reasoning system (besides generating memory; see sect. R1.3). I envisage that the latter is capable of only a very limited range of inferences (e.g., selecting the currently strongest desire among its inputs and using conditional inference; I presume that it has no capacity to *conjoin* inputs). And I envisage that it can only draw on a limited database of information (e.g., a set of abstract motor-schemas of performable actions and action sequences). So, the practical reasoning system, too, deserves to be regarded as a module (Carruthers, forthcoming b).

The thesis that language is the vehicle of intermodular integration is intended to apply *only* at the interface between the conceptual belief-forming and desire-forming modules and the practical reasoning system. The claim is that it is *those* contents (the outputs of the conceptual modules, fit to be taken singly as input by the practical reasoning system) that can only be conjoined with one another for purposes of further inference (in contrast to the purposes of storage or memory; see sect. R1.3) through the offices of the language faculty. The fact that modular outputs can be integrated at other levels of cognition is simply irrelevant.

The following puzzle arises, however. Given that modular outputs can be integrated at other levels in cognition, how is it that (prior to the evolution of the language faculty) this *wasn't* the case at the interface between conceptual modules and the practical reasoning system? The answer is that cognitive systems only evolve the powers that they need to, whereas cross-module talk is computationally expensive (**Bryson**); and it isn't clear that there was any *need* to integrate module-specific beliefs and desires.

Consider the powers that a practical reasoning module might have in the absence of language. (For a more elaborate and complex account, see Carruthers, forthcoming b.) It takes as input whatever is the currently strongest desire *P*. It then initiates a search for beliefs of the form $Q \rightarrow P$, cueing a search of memory for beliefs of this form and/or keying into action a suite of belief-forming modules to attempt to generate beliefs of this form. When it receives one, it checks its database to see if *Q* is something for which an existing motor-schema exists. If so, it initiates a search of the contents of current perception to see if the circumstances required to bring about *Q* are actual (i.e., to see not only whether *Q* is something doable, but doable here and now). If so, it goes ahead and does it. If not, it initiates a further search for beliefs of the form $R \rightarrow Q$, and so on. Perhaps the system also has a simple stopping rule: If you have to go more than *n* number of conditionals deep, stop, and move on to the next strongest desire.

In addition, one thing that immediately ancestral practical reasoning systems would have had, surely, is the power to initiate episodes of *mental rehearsal.* Once some sort of initial plan had been hit upon (I want *P;* if I do *Q*, I'll get *P; Q* is something I can do right now), there would have been considerable value in feeding the supposition that I do *Q* back through the various central modular systems once again as input, to see if such an action would have other as-yet-unforeseen consequences, whether beneficial or harmful. This doesn't seem to require anything amodular to be built into the system yet; for both the desire for *P* and the belief that if *Q* then *P* can be products of individual modules.[4,5]

The point is that earlier hominid species, armed with some such practical reasoning module and a bunch of powerful belief-forming and desire-generating modules, would already have been pretty smart. They could have generated flexible and appropriate plans of action, with the different elements of the plan drawn from different conceptual-modular domains. I thus agree with **MacWhinney** that our hominid ancestors would have had sophisticated planning abilities prior to the evolution of language. But in contrast with him, I don't think that they would have needed to be capable of integrating the outputs of conceptual modules for this to be possible. For it simply isn't clear why they would have *needed* to do cross-modular integration of contents. And on my account this only became possible as a by-product once a language faculty had evolved for purposes of communication, given that such a faculty would have been set up in such a way as to receive, and to combine, contents from each of the conceptual modules.

### *R1.3. Mental modularity and domain-integrating concepts*

As **Robbins** rightly notices, there is a sense in which my thesis is intended to apply only at the level of *process* or *inference,* not at the level of stored content. Indeed, modularism, too, is a thesis about processing and not about storage. Modules are learning mechanisms, charged with generating information and appropriate goals. Whether memory systems are also fractionated along modular lines, or whether their patterns of organization need to follow the *same* fault-lines, is quite another matter, as shown in this section. So the fact that there exist domain-integrating *concepts,* in the way that both **Hurford & Dessalles** and **Nelson** argue, poses no special threat to that thesis.

For example, it is plausible that prelinguistic representations of individual people should be organized in folders, with some sort of identifying label (perhaps a recognition schema), and with information about the individual drawn from a variety of modular domains (physical, biological, psychological, social, and so forth) contained within the folder, and so in one sense *conjoined.* But I claim that to *do* anything much *with* the information in a folder in the absence of language (to draw further inferences from it of a nonpractical sort), one would have to take each item of domain-specific information and process it separately within the appropriate conceptual module.

Similarly, then, for the prelinguistic event-concepts that combine information from various domains, discussed by **Nelson** I can allow their existence, consistently with my thesis, by claiming that the drawing of inferences from such concepts in the absence of language will always be a piecemeal and module-specific affair.

What exactly a modularist should say about any given domain-integrating concept will depend on the details of the case, of course. In addition to the possibility just canvassed in connection with concepts of individuals and of event-types, there are at least two further possibilities. Each of these can be illustrated through discussion of the abstract concept *three,* raised as an example against me by **Hurford & Dessalles**. Let me elaborate.

The concept *three* can, of course, be applied across all domains, wherever there are individuals to be numbered. So one can have three *rocks,* three *cats,* and three *thoughts.* But on some views that concept might still be proprietary to a more abstract numerical module (Butterworth 1999), whose operation is dependent on classifications effected by other modular systems.[6]

An alternative view would be that the abstract and exact numerical concept *three* is dependent on language (Spelke

& Tsivkin 2001). All that would be available in prelinguistic infants and monkeys, underpinning the behavior that seems to approximate to possession of that concept, would be the object-file system. This can hold in working memory a representation of three separate concrete individuals, enabling subjects to track the movements and absences of those individuals (Carey 2001).

### R1.4. Mental modularity and distinctively human thought

The central challenge for any strongly modularist account of human cognition, of course, is to account for the distinctive flexibility and creativity of the human mind, which seems to mark us off so sharply from all other species of creature on earth (past or present) in the way that **Bickerton** emphasizes. The main thesis of the target article is intended as a start on this problem: It is cross-modular contents formulated in language and cycles of processing activity involving language in "inner speech" that are supposed to account for some of that flexibility. But this was, of course, only a start and cannot by itself be the whole story. Let me briefly mention two other components.

One (not mentioned in the target article) is a supposition-generator or *supposer* (Nichols & Stich 2000), which in my view is built on the back of the language faculty, picking up on weak associations and analogies to generate novel sentences for consideration. It is this that is responsible for the distinctive creativity of human thinking, on my account. And I claim that it is the function of the species-specific disposition to engage in *pretend play* in childhood to build and hone the activity of this supposition-generator (Carruthers 2002b).

At this point I differ from **Bickerton** (and also, I think, from Chomsky 1988). He seems to believe that the human language-faculty by itself is sufficient to explain our disposition to engage in creative thought. But this can't be quite right. It is one thing to be *capable,* by virtue of the combinatorial powers of one's language, of entertaining novel and creative sentences. With this I agree. And it is quite another thing to have any *disposition* to *generate* such sentences. This is what the supposer confers on us, on my account.

The other major element in an account of human creative powers is some sort of faculty of abductive inference, or faculty of inference to the best explanation, that can receive these novel sentences and evaluate their worth. This was mentioned briefly in the target article, where I gave the misleading impression that I thought it would be embodied in a separate module. And to this **Dale & Spivey** rightly object that a faculty of abductive inference would seem to be the archetype of a radically *a*modular, or *holistic,* system (Fodor 1983).

In fact, my hope is that a kind of "virtual" faculty of abductive inference can be built out of elements already present in module-based cognition. For details, see Carruthers (forthcoming a). But very roughly, the idea is that principles of *simplicity, relevance,* and *coherence with existing belief,* which are already employed in the interpretation of speech and the evaluation of speech-dependent testimony from other people, would *become* a faculty of inference to the best explanation when directed inward, in cycles of "inner speech," giving rise to a preference for simple but fecund theories constrained by background belief.

The overall goal of the research project (of which the target article is one component) is to try to explain distinctively human cognition by seeing it as built out of modular components. And the motivation for adopting that goal, in turn, is to make it possible for us to see human cognitive processes as computationally realized, given that computational processes have to be encapsulated if they are to be tractable.

### R1.5. LF and logical inference

The hypothesis tentatively advanced in the target article is that it is linguistic representations of a certain specific kind that serve as the medium of intermodular integration. These are, namely, representations at the level of *logical form,* or LF (Chomsky 1995; Hornstein 1995). In this section I respond to a number of different objections to this idea. But first let me stress that such a claim isn't strictly essential to the main thesis of the target article, which is just that representations in a modular language faculty (whatever form they might take) serve to integrate contents across conceptual modules.

**Dale & Spivey** challenge the very existence of LF, or at least its psychological reality. But in doing so, they present a challenge to a significant portion of contemporary linguistics. I take my stand with the latter. For given Chomsky's current construal of LF, in which LF and PF (*phonological form*) are the *only* linguistic levels (Chomsky 1995), to deny LF is tantamount to denying abstract syntax.

In fact, evidence of the psychological reality of LF is legion. For example, not only adults but children as young as three years old will *only* accept the contraction of "want to" to "wanna" in those cases (statistically by far the most common) where there isn't a covert *wh*-trace between the verb and the preposition (Pietroski & Crain 2001; Thornton 1996). Thus subjects will accept the contraction of (1) to (2), but not of (3) to (4).

(1) I don't want to make breakfast for John.
(2) I don't wanna make breakfast for John.
(3) Who does John want to make breakfast?
(4) *Who does John wanna make breakfast?

This is because the LF representation of (3) contains a covert trace of the subject of "make breakfast," as in "$\text{Who}_t$ does John want *t* to make breakfast?"

Ironically, indeed, the very data cited by **Dale & Spivey** in their attack on LF, concerning the partial and incomplete nature of semantic processing (Ferreira et al. 2002; Sanford & Sturt 2002), are fully in accord with Chomsky's own thinking (Chomsky 2000; Pietroski 2003). According to Chomsky, the semantic contribution of the language faculty and of LF stops short of a full specification of truth-conditions. Thus the full LF representation of the sentence, "He saw the book in the window and bought it," would leave open the question of whether it was the very *token* book seen that was bought or just an instance of the same *type.* It is then supposed to be the work of other faculties of the mind to resolve such ambiguities. And presumably such resolutions are only sought when needed, just as Sanford and Sturt (2002) suggest.

**Molina** wonders why I chose LF to play the integrative role and doubts whether it is plausible to say that LF representations could constitute "the *locus* of thought (where thinking is really happening)." But, *of course,* I don't believe that *only* LF thoughts are real! I believe that the individual outputs of the conceptual modules, prior to integration in

LF, are also thoughts, and that thinking is going on within each of those modular systems and also in the practical reasoning module. LF plays a constitutive role in only a very limited range of thoughts, in my view, namely, those that combine together the outputs of conceptual modules for further processing.[7]

But why LF? Partly because my thesis is that it is the natural language faculty that does the integrative work. And partly because, according to our best theories, LF is the place where the language faculty interfaces with other systems in such a way as to fix external reference, and these systems would presumably include the various conceptual modules whose outputs stand in need of integration.

**Chater** wonders whether there is really any difference between my views concerning the role of LF and the traditional thesis that there is a *language of thought* (LOT, or mentalese) over which logical and other forms of inference are defined. And he wonders whether the difference between my view and that of, say, Fodor (1975) might be merely terminological.

One difference is that LF representations will contain lexical items drawn from a specific natural language (and might also contain some language-specific structural elements; see the discussion in sect. R5.6), whereas LOT sentences are supposed to be in some sort of universal "brain language." Put differently, the point is that the LF representations of a speaker of English will contain lexical items drawn from that very language; whereas LOT sentences won't (in general) do so (except in the case where a subject is thinking *about* English sentences).

Another difference between LF and LOT is that I certainly *don't* envisage that there is any kind of global logical capability operating over sentences of LF. A language module should, of course, be capable of certain *limited* forms of inference defined over LFs (see later in this section). But such inferences need not be manifestations of any sort of general-purpose inference engine. In fact, I very much doubt whether humans *have* a global logical faculty (Evans & Over 1996; Gigerenzer & Todd 2000). Rather, a set of heuristics and implicit procedures (together with the limited inferential powers of the language faculty) could collectively provide a sufficiently good *approximation of* logical ability.

Of course, what we do also have, if dual process theories of cognition are correct (Evans & Over 1996; Frankish, forthcoming; Stanovich 1999; and see sect. R.4.4 here), is the capacity to form explicit beliefs about which forms of inference are and aren't valid and to train ourselves to follow only valid patterns in our explicit thinking. But this capacity, in my view (although also involving natural language sentences as vehicles of thought), supervenes the more basic role of LF in integrating domain-specific thoughts.

**Robbins** correctly notices that the integrative role of the language module, on my account, depends on it having the capacity for certain kinds of inference. Specifically, it must be capable of taking two LF sentences, constructed from the outputs of two distinct conceptual modules, and combining them appropriately into a single LF representation. He claims that such a view is highly implausible, however, and that it puts me into conflict with the views of most contemporary linguists. But, this is not so.

**Robbins**, like **Chater**, mistakenly thinks that what is in question is the existence of a general-purpose inference engine located within the language faculty. For, indeed, such a claim is not only intrinsically unbelievable, but certainly wouldn't be believed by any working linguist. However, just about every linguist *does* think (with the possible exception of Fodor and those closely influenced by Fodor) that *some* inferences are valid, and are known to be valid, purely by virtue of our linguistic competence.

Most linguists believe, for example, that the inference from "John ran quickly" to "John ran" is endorsed by virtue of semantic competence (Parsons 1990); and many would claim that such competence is embodied in the language faculty. In contemporary parlance, this amounts to saying that the inference is made by the language module transforming sentences of LF. Indeed, it is worth noting that one way in which such inferences might be handled is in terms of covert quantification over events and conjunction reduction (Pietroski 2004). Thus on this sort of account, "John ran quickly" really has the form, ∃*e (e is a running & e is by John & e is quick).* And then the inference to "John ran" is just an instance of conjunction elimination.

In conclusion, then, so far from being implausible, the idea that certain limited forms of inference (notably conjunction introduction and conjunction elimination, among others) are handled internally within the language faculty in transformations of LF sentences is pretty much the orthodox view in contemporary linguistics.

## R2. Mechanisms of modular integration

In this section, three distinct proposals made by commentators will be discussed concerning the mechanisms for integrating central-modular domain-specific contents.

### *R2.1. The relative pronoun*

**Bermudez** appears to endorse both the spirit and the letter of my proposal but offers a friendly elaboration. Instead of saying merely that it is sentences of LF that serve to integrate modular outputs, why not say that it is the availability of the relative pronoun, in particular, which makes this possible? I would be happy to accept such an amendment if I thought that it was well motivated. But I can't see that it is.

Consider the simplest possible example arising from the Hermer-Vazquez et al. (1999) re-orientation experiments: Suppose that the sentences to be conjoined are "The toy is near a long wall," and "The toy is near the blue wall" (in a case where the blue wall is also a long wall). My own proposal is just that, with the referent of "wall" indexed as the same, LF structures entitle us to multiply embed the predicates "long" and "blue" to form the sentence, "The toy is near the long blue wall."

**Bermudez**'s proposal amounts to the claim that the first of the preceding sentences really has the form, "The toy is near a wall *that is long*," and the second has the form, "The toy is near the wall *that is blue*"; and then (because the referents of "wall" are indexed as the same) we are entitled to move to the single sentence, "The toy is near the wall *that is long and that is blue.*" In effect, the proposal seems to be that all complex predicates are relative clauses. But I doubt whether many linguists would consider such a suggestion worth pursuing. For what could be wrong with allowing "blue" and "long" to attach to a noun phrase directly, without the formation of a relative clause?

**Bermudez**'s worry (on my behalf) would seem to be this: If the predicates "blue" and "long" express concepts drawn from different modular domains, then there is nothing to guarantee that it will even make *sense* to attach them to objects drawn from the other domain. And the abstract apparatus of relative clauses is somehow supposed to help with this. I don't really see how the story here is supposed to go, in fact. But in any case, the original worry is ill motivated and betrays a misunderstanding of the sense in which conceptual modules are domain-specific.

Many modules can process information concerning the same object or situation, and so far as I can see, this will be the general case. But then we automatically have a guarantee that predicates produced by the one module will apply to the objects of the other. This should have been manifest from the original example. The thoughts produced by the geometry module and the landmark module *both* concern the location of the toy; and it is the toy that is the subject of both resulting sentences. Equally, one and the same wall can be *both* long (spatial module) *and* colored (color module).

The problem to which the language faculty provides a solution, in fact, *isn't* that of generating abstract concepts that can apply across modular boundaries (for this already comes for free with the modules themselves); it is rather to provide an abstract framework within which concepts of the different kinds can be combined into a single thought for the purposes of inference.

### R2.2. Metaphor and structure mapping

**Bryant & Gibbs** and **Dominey** each propose that the human cognitive system contains a prelinguistic capacity to map one domain onto another, giving rise to metaphorical conceptual schemes. For example, the domain of debate and argument can be mapped onto the domain of warfare, giving rise to an elaborate system of concepts summarized by the slogan, "Argument is war" (Lakoff & Johnson 1980). And it is suggested that such conceptual schemes are nonlinguistic in character, receiving expression *in* metaphorical language rather than being reflections *of* it. So on such an account, the basic mechanism of intermodular integration will be this domain-general, structure-mapping capability rather than language.

The fact that there exists a correlation between cross-domain mapping and metaphorical language is one thing, however; the question of the direction of causation here is quite another. It might be possible to claim that such conceptual mappings begin initially with the integrative functions of language. Or so, at least, I shall argue.

For these purposes, it is important to distinguish between live metaphors and dead ones. Live metaphors are fresh creations and strike us (if they are effective) with the force of insight; dead metaphors are once-live metaphors that have been used so commonly that it takes hardly any effort to process them, and their metaphorical status is barely discerned. For notice that the metaphorical conceptual systems studied by Lakoff and Johnson (1980; 1999) and others are dead ones, enshrined in common patterns of speech. If someone says, for example, "Don't worry; I'm on top of the situation," we probably won't even notice that a metaphor has been used.

Recall from section R1.4 that my account presupposes the existence of a language-dependent *supposer.* It might work somewhat as follows. Sometimes when two previously unconnected concepts *A* and *B* are coactivated (albeit one of them weakly) by the same stimulus, situation, or thought, then this leads to the supposition that *A is B* being explicitly formulated. This supposition is then fed as input to the various (mostly modular) reasoning systems, and the results evaluated. At the outset of an episode of pretend play in childhood, for example, the sight of a curved banana might also activate the concept *telephone* (because of the banana's similarity in shape to a telephone handset). The child then entertains the supposition, "The banana is a telephone." From this supposition various inferences can be drawn. If the banana is a telephone, then you can call grandma on it, and it will have buttons down its middle section for dialing. The child can then access these consequences, pretend to dial, and subsequently begin a pretend conversation with grandma.[8]

Generalizing from this example, then, one might propose that a live metaphor is first created when some sort of structural correspondence or similarity between two phenomena leads to at least weak coactivation of the concepts applying to each. The resulting metaphorical thought is formulated in language by the suppositional system and fed back to the various modular inferential systems. Consequences that are evaluated as true or useful are retained, and the rest discarded. (For example, the playing child is likely to discard the inference, "If the banana is a telephone, then it won't squish if I squeeze it.") This account presupposes no domain-integrating capacity beyond the language-dependent supposer.

In terms of such an account, it is easy enough to see how culturally entrenched dead metaphors might become established. Wherever two domains are structurally alike in some significant respect, this is likely to lead to weak coactivation of the corresponding concepts, and hence people will sometimes entertain the supposition that the one domain (or an aspect of the one domain) *is* the other. If there are enough structural similarities between the two domains for the one to serve as a viable partial model of the other (as in the case of "argument is war"), then the corresponding modes of speech are likely to become entrenched. Hence we now talk of "winning an argument" and "attacking someone's premises" without giving it a second thought.

In conclusion, then, I claim that one can explain the deeply metaphorical character of much of our language and speech consistently with my main hypothesis, without having to suppose that people possess some sort of language-independent, domain-general, structure-mapping capability.

### R2.3. Metamodule or language module?

**Atran** presents the following important challenge to my account: Why should we think that it is language that does intermodular integration rather than some sort of metarepresentational faculty (or a theory of mind mechanism [ToMM])? Atran agrees with the modularist framework adopted in my paper; and he agrees that domain-general, flexible cognition somehow has to be built out of modular components. But he sees nothing yet to discriminate between my proposal and the idea previously advanced by Sperber (1996) that ToMM is the module that has the power to combine together the outputs of all the others.

This alternative proposal is worth replying to at some length. So consider, then, the competitor hypothesis that it is ToMM that has access to the outputs of all the various central or conceptual modules, and that also has the power to combine those contents together to form intermodular thoughts.

One difficulty for such an hypothesis is that there is evidence that ToMM doesn't, in fact, have any such access. Instead of automatically receiving as input the various beliefs that influence our own behavior (which would then make self-explanation relatively trivial), there is considerable evidence that what ToMM does instead is engage in *self-interpretation,* ascribing thoughts and goals to oneself to explain manifest behavior in something very much like the way that thoughts and goals are ascribed to other people (Gazzaniga 1998; Gopnik 1993).

This emerges most clearly in those many cases where it can be demonstrated that people's attributions of thoughts to themselves are really instances of *confabulation,* rather than mere reporting of contents to which ToMM has some sort of direct access (Gazzaniga 1998; Gopnik 1993; Nisbett & Ross 1980; Nisbett & Wilson 1977; Wilson 1985; 2002; Wilson & Stone 1985; Wilson et al. 1981). Indeed, a strong case can be made that the only circumstances in which ToMM has non-inferential access to a subject's own thoughts is where those thoughts are linguistically expressed in inner speech (Carruthers 1998b; see also the discussion in sect. R5.2).

In conclusion, if ToMM doesn't have *access* to the subject's beliefs (i.e., the outputs of the subject's belief-forming conceptual modules), then it cannot, of course, be ToMM that serves to bind together and integrate the contents of those beliefs on a regular basis. In contrast, we know that the language faculty must have access to (a significant subset of) the subject's beliefs for those beliefs to be expressible in speech.

Another difficulty for the Sperber-**Atran** proposal (even if we set aside the first) is to explain why any such metarepresentational intermodular architecture should have evolved. For the main business of ToMM is presumably to predict and explain the behavior of conspecifics. This capacity would only require the construction of intermodular thoughts if others were *already* entertaining such thoughts. Otherwise, attributions of domain-specific thoughts alone would do perfectly well. By contrast, in the case of language, the demands of swift and efficient communication would have created a significant selection pressure for intermodular integration, allowing the outputs of distinct central modules concerning the same object or event to be combined into a single spoken sentence. So instead of saying separately, "The object is near a long wall," and "The object is near the blue wall," one can say much more succinctly, "The object is near the long blue wall."

It might be replied that the pressure for ToMM to integrate contents across modules could have come, not from the demands of predicting and explaining the behavior of oneself and others, but rather from the benefits that such integration can bring for other areas of activity (such as solving the reorientation problem). This is possible. After all, it is common enough in biology for a system initially selected for one purpose to be co-opted and used for another. And it might be claimed that ToMM would be ideally placed to play the integrative role, taking inputs from all the other central modules (allegedly anyway; see earlier in this section) and providing outputs to practical reasoning.

Actually, it is hard to see how one could get from here to anything resembling the full flexibility of human cognition. For there is no reason to suppose that ToMM would have been set up so as to provide its output *as input to* the other central modules, in which case there would be no scope for cycles of reasoning activity, with ToMM combining outputs from central modules and then harnessing their resources for further reasoning. In contrast, because language is both an output *and* an input module, it is well positioned for just this role, as I argued at length in the target article.

What of the experimental evidence? **Atran** attributes to me the claim that dual-task data of the sort collected by Hermer-Vazquez et al. (1999) cannot discriminate between the idea that it is language, on the one hand, or ToMM, on the other, that performs the integrative function across modules. But this is a misreading of the target article. What I said, rather, is that dual-task studies cannot easily be used to test whether it is language that plays the role of integrating the outputs of *ToMM* and *other modular systems,* because of the likelihood that ToMM itself sometimes accesses and deploys the representations of the language faculty in the course of its normal operation. For then any task that disrupts language is likely to disrupt ToMM as well. But this is by no means the same as saying that dual-task studies can't discriminate between ToMM and language as candidates for playing the integrative role in respect of *other* modular systems (i.e., systems other than language and ToMM).

Indeed, the dual-task data collected by Spelke and colleagues seem to me to support my own proposal quite strongly, when matched against the Sperber-**Atran** one. For if subjects fail at the task in the speech-shadowing condition (but not in the rhythm-shadowing condition) because they cannot then integrate geometrical and landmark information, it is very hard to see how it can be a disruption to ToMM that is responsible for this failure. For there is no reason to think that shadowing of speech should involve the resources of ToMM.

Admittedly, a good case can be made for saying that normal speech comprehension and production will implicate a subsystem of ToMM, at least, where that subsystem is charged with figuring out the conditions for *relevance* in communication (Sperber & Wilson 2002). So normal speaking and comprehending would surely disrupt any other ToMM-involving task (and so any task requiring intermodular integration, on the hypothesis that it is ToMM that performs this role). But there is no reason to think that this should be true of speech *shadowing.* For in this case there is no requirement of comprehension and no attempt at communication (see also sect. R3.2).

In conclusion, then, I claim (on theoretical, evolutionary, and experimental grounds) that language is a much more plausible candidate for integrating the contents of other central modules than is ToMM.

## R3. The reorientation data reconsidered

Several commentators raise doubts about the reorientation data collected by Spelke and her colleagues (Hermer & Spelke 1996; Hermer-Vazquez et al. 1999). Responding to these doubts is the task of the present section.

### *R3.1. Language as an aid to memory*

**Bryson** and **Clark** each provide very similar alternative explanations of the Spelke data, albeit from quite different theoretical perspectives (**Bryson** is a modularist; **Clark**, a connectionist). Their idea is that linguistic terms might serve, not as vehicles of intermodular integration, but rather as aids to memory, helping us to focus our attention on tasks that we would otherwise find difficult.

What this suggestion overlooks, however, is the datum that the only predictor of success in the reorientation tasks that could be discovered among young children was productive use of "left" and "right" (Hermer & Spelke 1996). For we already know that prelinguistic children and rats rely on geometric information in attempting to solve these tasks. So their problem is *not* to recover the spatial relationships involved when they have been disoriented: We know that they can already make use of those. Rather, their problem is to combine spatial information with landmark (e.g., color) information.

Because it seems to be colors that children have trouble remembering in these circumstances, a memory-enhancement theory should predict that it would be the productive use of color vocabulary that would be the best indicator of success. But this isn't what we find. What enables children eventually to succeed in the task isn't some way of remembering the relevance of color (memory enhancement), but rather a way of combining into a single explicit goal *both* items of relevant information (i.e., module integration).

But then why, on the module-integrating account, is it spatial vocabulary that is crucial? Shouldn't color vocabulary be just as important? Indeed so; but the fact is that color vocabulary is much more easily acquired. So the crucial watershed, at which *both* types of lexical item become available together for the first time, is the acquisition of "left" and "right."

**Wynn & Coolidge** put forward a related counterproposal for explaining the Spelke data. Relying on Baddeley's (1986) model of the working-memory system as being made up of a central executive and two slave systems (the phonological loop and the visuo-spatial sketchpad), they suggest that the reason why performance is disrupted in the speech-shadowing task might be because of interference with the central executive component of working memory. And because the central executive isn't supposed to be language dependent, this means that integration of contents across domains shouldn't be dependent on language.

Without a lot more detail, however, this proposal doesn't look well motivated. For Hermer-Vazquez et al. (1999) go to great lengths to argue that the rhythm-shadowing task is just as demanding of the resources of working memory as the speech-shadowing one. So **Wynn & Coolidge** need to show us how there is something about the speech-shadowing task that is especially demanding of the resources of the central executive in particular.

### *R3.2. A contrast otherwise explained?*

Samuels raises an objection that, if it were successful, might provide just the materials that **Wynn & Coolidge** need.[9] He points out that on orthodox theories of speech production, two additional kinds of cognitive resource are involved, neither of which would be required for the rhythm-shadowing task. First, extralinguistic concepts have to be recruited and combined together into a representation of the message to be communicated; and second, communicative intentions have to be formed. In that case, speech-shadowing might differentially disrupt performance because forming communicative intentions places greater demands on working memory than do mere motor intentions, and/or because the resources used for combining concepts are otherwise occupied by shadowing, not because language per se is the medium of intermodular integration.

Unfortunately for **Samuels**'s proposal, however, there is no reason to think that speech *shadowing* uses the same cognitive resources as does normal speech *production,* and there is good reason to think that it doesn't. First, when shadowing speech, people don't have any intention to communicate, so no communicative intentions need to be formed. And second, because there isn't any message to be communicated either, there is no reason to think that extralinguistic concepts would have to be combined together with one another.

It looks as if speech shadowing should only need to use some sort of fast-track loop between phoneme perception and phoneme production, indeed, bypassing the conceptual systems altogether. And the *speed* of speech shadowing might seem to suggest that this is in fact the case. Speech can be shadowed with latencies as small as 250 milliseconds, with subjects beginning to pronounce each phoneme as soon as it has been perceived. Indeed, these were among the data used by Fodor (1983) to argue for the modularity of the language faculty, suggesting that such shadowing can be carried out independently of conceptual thought.

Admittedly, there is evidence from instances of speech repair that shadowers are sometimes influenced by syntactic and semantic context, in cases where they are shadowing speech that contains errors (or at any rate, this occurs when they are concentrating exclusively on the shadowing task; see Marslen-Wilson 1975).[10] But given the speed with which repair is conducted, it seems highly likely that the syntactic and logical forms implicated are all of the "narrow" sort that are internal to the language faculty (see sect. R1.5). In that case, it is unlikely that conceptual systems external to the language faculty of the sort that (on the orthodox account) would be needed to solve the reorientation task are being accessed.

### *R3.3. Do fish integrate across modules?*

**Vallortigara & Sovrano** remind us of the fact that a range of species other than humans and rats (specifically, monkeys, chickens, and fish) are able to solve the reorientation task in the absence of language. Many researchers have drawn from this fact the conclusion that integration of information across the relevant modules can't, therefore, be dependent on language.

In the target article I pointed out that this conclusion is too hasty, because *sequential* use of domain-specific information would also be sufficient to solve the task. I suggested that the practical reasoning system in fish, say, might be set up in such a way as to embody the rule, "When disorientated, look for landmark information first [e.g., *red wall*], and then geometric information second [e.g., *long wall on left*]." This would generate the correct solution. In human children and rats, in contrast, the rule would appear to be "When disorientated, look for geometric information, period."

**Vallortigara & Sovrano** acknowledge that this is an abstract possibility; but they complain that it doesn't appear to be empirically testable. And they conclude that in such circumstances my proposal would be a "supposed difference [that] would not make any difference." But this is far too strong. Even if one thought that real theoretical differences have to be testable in principle, it is quite another matter to insist that they be testable *now.* And scientists frequently articulate genuine theoretical contrasts that no one can see how to test at the time.

What I do claim on behalf of my sequential-ordering proposal is that it is theoretically well motivated. First, task-analysis of what a practical reasoning system has to be able to do suggests just such a picture (articulated briefly in sect. R1.2). Given a goal, the practical reasoning system has to be able to call up a sequence of conditional beliefs that will guide the organism toward achieving that goal.

Second, consider the significance of the fact that human children and rats are perfectly capable of using both geometric and landmark information in other circumstances. If the use of two forms of information in the same task means that they are routinely *integrated,* in the way that **Vallortigara & Sovrano**'s interpretation of their data implies, then it would be puzzling why such integration should suddenly become impossible under conditions of disorientation. If we suppose that the information is only drawn on sequentially, however, then it makes perfectly good sense that some organisms might have an innate preference for just one of these kinds of information when disoriented.

### *R3.4. Success in a different task*

**Varley & Siegal** describe a piloted experiment in which aphasic patients were tested on a version of the task just discussed, but with a scale model of a rectangular disorientation room being spun in front of the patient, rather than the patient being disoriented *in* such a room. And they report that all subjects were capable of solving the task, using both geometric and landmark information in the absence of language.

Unfortunately, this is just a different task, and no conclusions can be drawn. There is no reason to think that subjects become disoriented in such circumstances. And there are no data from rats or young children to suggest that the task should be a difficult one. As I remarked in section R3.3, we already know that rats and young children without language can use both geometric and landmark information when not disoriented; and I have just argued that sequential use of such information might be what enables fish and chickens to solve the reorientation tasks. That may be how the aphasic patients solve the spinning-model tasks, too.

## R4. Competing claims for the involvement of language in thought

A number of commentators advance claims for the positive involvement of language in thought that are inconsistent with the main theoretical idea (central-module integration) defended in the target article. These claims will be discussed in this section.

### *R4.1. No thought without language?*

**Gauker** defends the traditional philosophical thesis that there can be no conceptual thought without language. (This also seems to be tentatively endorsed by **de Villiers & de Villiers** as a generalization of their claims concerning higher-order thought; see sect. R4.3) I claim, on the contrary, that there can be a great deal of conceptual thought in the absence of language, namely, the outputs of the conceptual modules and the inputs to practical reasoning.

**Gauker** proceeds by defining conceptual thought to mean: having the capacity to represent a particular thing as belonging to some general kind. He then tries to explain away the behaviors that might seem to suggest that nonlinguistic animals have any such capacity (e.g., by appealing to "elaborate mental movie making"). There is nothing wrong with the definition; but the attempts at explaining away limp badly, in my view.[11]

**Gauker** tries to explain animal capacities for categorization in terms of an ability to discern similarities amongst perceptual images. But this surely isn't sufficient. All prey species have the capacity to distinguish between *predator* and *nonpredator,* for example. But if all such judgments were made in terms of mere perceptual similarity, one would expect warthogs and hyenas to be put together in one category and lions in another.

Many social species (e.g., monkeys and apes) can recognize individuals and are capable of tracking the changing properties of those individuals. They behave appropriately toward an individual depending on what happened at their last encounter or depending on observed one-off encounters with others. Why this shouldn't count as "representing a particular thing as belonging to a general kind" escapes me. And I have no idea how one might account for such abilities in terms of sequences of sensory images, without this tacitly amounting to the claim that the animals in question engage in conceptual thought.[12]

### *R4.2. No off-line thought without language?*

In the target article I interpret Bickerton (1995) to be claiming that all imaginative thinking that doesn't simply reflect current or remembered circumstances requires language. In his comment **Bickerton** corrects me, saying that he had meant that language is the vehicle of genuinely *creative* thought, including the capacity to conceive of truly novel objects or circumstances. With such a claim I can agree. But I deny that language is *sufficient* for us actually to entertain such thoughts. Rather, as pointed out in section R1.4, I think that we need to postulate a distinct language-dependent *supposer,* whose activities are honed in childhood through the disposition to engage in pretend play (Carruthers 2002b).

**Bickerton** goes on to suggest that the capacities to entertain conditional thoughts of the form, "If *x* then *y,*" and to entertain causal thoughts of the form, "*x* because *y,*" are dependent on language. With this I strongly disagree. It will be the business of the conceptual modules (existing prior to and independently of language) to generate such thoughts. For example, the folk mechanics module can surely generate thoughts of the form, "The rock striking the core caused the sharp edge." And such thoughts would have been available to hominids long before the evolution of language. (Indeed, rats, too, seem to be capable of causal thinking; see Dickinson & Shanks 1995.)

On the model sketched in section R1.2, similarly, it is the main business of *all* belief-forming modules to generate conditional information for input to the practical reasoning system. For example, a prelinguistic hominid engaged in a hunt might have needed some way to attract the attention of a cohunter without alerting the prey. The theory of mind module (or a precursor thereof) might have generated the conditional, "If a pebble were to drop just in front of him, it would attract his attention," hence yielding an appropriate plan (to throw a pebble). I see no reason to assume that such routine planning should require language.

### R4.3. No higher-order thought without language?

**De Villiers & de Villiers** present their claim that language is the vehicle of higher-order thought, or thoughts about the thoughts (e.g., the false beliefs) of another. I allow one aspect of this claim. Specifically, I allow that thought about *domain-integrating thoughts* is dependent on language. For if the very existence of such thoughts is dependent on language, as I argue, then thoughts *about* them would have to involve language, too. But I deny that language is necessary for higher-order thought in general.

I make this denial, both because I think it very plausible that the evolution of the language faculty would have presupposed the prior existence of a capacity for higher-order thought (Origgi & Sperber 2000), and because I think that the experimental evidence supports it. Let me elaborate a little on the latter point.

As **de Villiers & de Villiers** note, I appeal for support to the experimental finding of Varley (1998), that a man with severe agrammatic aphasia (and no mentalistic vocabulary) was nevertheless easily able to solve the usual battery of false-belief tasks (explained to him by pantomime; itself an indicator of his sophistication in this domain). **De Villiers & de Villiers** point out the possibility that he might have deficits only of comprehension and production, however, leaving his underlying knowledge of LF intact. But, in fact, his input and output deficits match one another very precisely, suggesting a deficit in underlying competence. And, indeed, brain scans show that he has lost the use of almost his entire left hemisphere!

The data that **de Villiers & de Villiers** cite in their own support are that competence with verbs of communication (e.g., "says that") and verbs of thought (e.g., "believes that") predicts false-belief performance. But as Hale and Tager-Flusberg (2003) point out, these data are confounded. That competence with verbs of thought predicts competence with false-belief *might* be because the requisite mental concepts need to be in place first to understand those verbs, rather than because those verbs are constitutive of thought about thought.

Hale and Tager-Flusberg (2003) undertook their own training study, coaching preschoolers on verbs of communication and on false belief. They found that training on sentential complements did indeed boost false belief performance, just as **de Villiers & de Villiers** would predict. But they also found that training on false belief tasks boosted performance, too, yet had no reciprocal influence on language. This result counts strongly against their view.

Another recent study, by Woolfe et al. (2002), also counts against the view that natural language syntax is crucially implicated in mental state representation. It found that late-signing deaf children were significantly weaker on a range of theory of mind tasks, even when matched for syntactic ability with a group of early-signing (and younger) deaf children. (The two groups couldn't be matched for ability with sentential complements, because British Sign Language *has no* sentential-complement construction!) Woolfe et al. argue that exposure to and engagement in *conversation* serves to boost theory of mind ability, rather than language *per se* being directly implicated in theory of mind reasoning.

### R4.4. Dual-process theories and language

**Evans & Over** discuss the relationship between the main ideas of the target article and so-called dual process theories of the sort developed by themselves (1996), Reber (1993), and Stanovich (1999). (Dual process theories are also endorsed by **Frankish** and **Nelson**.) They find several deep connections between the two approaches.

They object, however, to the modularism that forms an important part of my approach. Although they allow that implicit "System 1" processes are generally domain-specific, they assert that much of this sort of cognition results from domain-general learning. But there is no reason why a modularist should insist that all modules are innate, in fact. We can allow that some modules are built in the course of development (Karmiloff-Smith 1992). This will be correct if chess-playing skill counts as modular, for example (possible), or if reading ability is so (very likely).

It is one thing to say that many implicit processes result from learning, however (and I agree); to say that they result from *domain-general* learning is quite another. For some modules might be built through the operations of other learning modules. And there might not really be any such thing as "general learning," anyway. For example, a strong case has now been made that even classical conditioning is best explained not as resulting from general associationist processes, but rather as arising from the computations carried by a specialized and innately structured foraging module, which calculates rates of return from different activities (Gallistel 2000; Gallistel & Gibson 2001).

Moreover, although **Evans & Over** accept that explicit System 2 thinking is closely related to language, they doubt whether it is conducted *in* language, citing evidence for the involvement of mental models in System 2 thinking. But they forget that natural language sentences form only one *part* of the overall process of explicit domain-integrative thinking, on my account. And they forget that mental models play an important element in my own views, enabling the natural language sentences that are being looped back in inner speech to be cast in a form accessible to the various conceptual modules once again.

I conclude, therefore, that the correspondence between dual-process theories and the main claim defended in the target article (the module-integrating role of language) may be much closer than **Evans & Over** suggest.

### R4.5. Language and connectionism

**Clark** and **O'Brien & Opie** put forward accounts of the role of natural language in cognition from a connectionist perspective.[13] Both comments focus on the role of inner speech in enabling different parts of the brain to communicate with one another and serving as a locus of attention and memory.

This might be all well and good if connectionism were viable as an overall model of cognition. But it isn't, for the reasons sketched in section R1.1. The successes of distributed connectionist models have been restricted to various forms of pattern recognition and feature extraction. And insofar as connectionist nets can approximate to anything resembling systematic thought and one-off learning, it is only through the imposition of an architecture that is tantamount to symbol processing (Marcus 2001).

It might be suggested that it is interesting, at least, to see that my thesis that language has an important cognitive function can have a close analogue within such a different theoretical framework. Well, yes, to a degree. But it isn't news that connectionism can find a place for natural language in cognition. Such views have been proposed since the early days of the movement (Rumelhart & McClelland 1986). And this is perhaps no accident, because natural language structures provide connectionists with their only means of approaching the serial and structured nature of conceptual thought. What *is* news is that language can be accorded a central role in the human mind within a framework that is realist and representationalist about structured thought and that is committed to a strongly modularist account of our overall mental architecture.

### *R4.6. Language as an invention for encoding platonic meanings*

**Baumeister & Vohs** deny that language evolved. Rather, they claim, it was invented. But this ignores the very significant body of evidence built up over the last half-century for the thesis that humans have an innately structured natural language faculty (for reviews, see Chomsky 1988; Laurence & Margolis 2001; Pietroski & Crain 2001). Some such view was taken for granted in the target article, and I stand by that.

According to **Baumeister & Vohs**, too, language is a tool for processing meanings, where meanings are abstract, mind-independent, platonic entities, discovered by humans when their brains reached a certain level of sophistication. In effect, this is just Popper's (1959) "third realm" doctrine all over again: (1) there is the physical world of objects and physical processes, (2) there is the mental world of experiences and ideas, and (3) there is the abstract world of propositions and meanings. The mental world is supposed somehow to "grasp" and make use of the abstract world in representing the physical one (see also Frege 1918).

The consensus among naturalistically inclined philosophers, however (once again for much of the last half-century), is that Platonism of this sort places insuperable obstacles in the way of achieving a scientific understanding of mind (Papineau 1993). We have reasonable hopes of understanding how minds can represent the physical world by being realized in computational processes in brains that engage in the right sorts of causal commerce with that world (Rey 1997). And abstract meanings might surely supervene upon such processes. But we have no hope at all of understanding, in causal and scientifically acceptable terms, how minds or brains might succeed in grasping elements of an abstract mind-independent reality ("meanings") directly, through which they would then succeed in representing the physical world.

## R5. Supplementary claims for the involvement of language in thought

A number of commentators present proposals that are supplementary to, and/or largely consistent with, the main theoretical idea (central-module integration) defended in the target article. These proposals will be discussed in this section.

### *R5.1. Language and concept formation*

Both **Hampton** and **Henser** claim that natural language lexical items serve as the vehicles of our more abstract concepts, serving to chunk otherwise unwieldy amounts of data into discrete and manipulable packages. This proposal strikes me as plausible, to a degree, especially in connection with concepts that have emerged out of many years of collective labor by scientists, and that can only be acquired by immersion in the relevant scientific theories, such as *electron, neutrino,* and *DNA.*

If **Hampton** and **Henser** intend their thesis to be very much stronger than this, however, then it seems to me almost certain to be wrong, because of the phenomenon of "fast mapping" (Bloom 2000). Children acquire lexical items at such a prodigious rate that the only plausible account of how they do it is that they are mapping those items onto previously existing nonlinguistic concepts, rather than constructing a set of language-involving concepts from scratch.

**Xu** introduces a fascinating body of data that suggest that natural language kind-terms are "essence place-holders." As I understand it, the thesis is exclusively diachronic: There is no suggestion that lexical items are constitutive of possession of the corresponding concepts. Nor does the thesis imply that language learning is the manner in which children acquire many new concepts, I think. Rather, use of kind-terms is taken by children to be *evidence of* the existence of an underlying essence to the kind. So the use of kind-terms is what enables children to select from among the wider set of concepts available to them the subset to which they are going to attach a belief in an underlying essence.

That children should possess some such disposition as this makes good sense, on the assumption that the language surrounding the child will embody the hard-earned wisdom of the society as a whole, concerning those kinds in the environment (especially biological kinds) that can underpin robust forms of inductive generalization. So in this sense, it may be that to learn a language is to tap into the beliefs of a culture (see also sect. R5.7).

### *R5.2. Conscious thinking and "inner speech"*

Both **Chater** and **Clark** are happy to accede to one of the subsidiary claims defended in my target article, namely, that "inner speech" is constitutive of *conscious* propositional thinking. But each also puts forward a stronger thesis, namely, that this is the *only* way in which we can entertain conscious propositional thoughts. Now as it happens, this is a thesis I am inclined to endorse and have defended elsewhere (Carruthers 1998b). But it requires a bit more work to establish than either **Chater** or **Clark** seem to allow.

The problem is that, besides reporting the occurrence of inner speech, many people also report engaging in pure wordless (nonlinguistic, nonimagistic) conscious thinking (Hurlburt 1990; 1993). So it needs to be argued somehow that these reports are illusory. And as it turns out, the confabulation data mentioned in section R2.3, which show that humans lack direct access to (most of) their thought processes, can be deployed to plug this gap (Carruthers 1998b); but it does need to be plugged, to get the conclusion that **Chater** and **Clark** both want.

**Frankish**, too, is happy to allow my claim that linguistically mediated thinking occurs consciously, in inner speech. But he challenges my suggestion that such thinking might also occur *non*consciously, through the manipulation of LF representations alone, without any phonological component.

I grant **Frankish** that the evidence of nonconscious linguistic thinking (e.g., the occurrence of "eureka thoughts") is far from conclusive. And the thesis isn't one that I need to place any weight on. But I still think that a case of sorts can be made. In particular, the well-known phenomenon of "sleeping on it" seems to me to support such a view. As we all know, when stumped by some difficult (conscious and linguistically formulated) problem, it often helps to sleep on it, or to turn our conscious thinking to other matters. And as a result, the solution sometimes just "comes" to us on waking, or in the midst of thinking about something else. These seem to be cases in which we can be confident that no conscious thinking about the problem has been going on in the interim (*contra* **Frankish**'s suggestion), which would imply that the thinking in question has been *non*conscious.

**Frankish** also argues that it is implausible to postulate *two* feedback loops: a phonological one, for conscious linguistic thinking; and an LF one, for nonconscious linguistic thinking (citing the work of Levelt 1989 in his support). But, actually, Levelt himself is committed to the existence of two feedback loops: a phonological one, for on-line speech repair; and a "message to be communicated" one, for monitoring one's own speech intentions. So it is open to me to claim either that this second loop runs on LF representations, or (more plausibly) that it can be exploited by LF representations in such a way as to subserve nonconscious module-integrating thinking.

In contrast, both **Bickerton** (in passing) and **Slezak** (at length) seem to challenge the very coherence of the idea that some thought might be conducted in inner speech. **Bickerton** offers no arguments. But **Slezak** claims that I have been seduced by the image of "hearing with the mind's ear" (cf. "seeing with the mind's eye"). He claims that my position commits me to endorsing one particular side of the "imagery debate" (Kosslyn 1994; Pylyshyn 2002), and that my account ends in incoherence as a result. There is much that goes awry here.

First, there is nothing fundamentally incoherent in Kosslyn's (1994) position on images. The claim that images are like percepts (indeed, perhaps *are* percepts of a self-induced sort), occupying the same cognitive resources as does perception, is a perfectly sensible (albeit controversial) one. And nothing about it need commit one to the idea of a little homunculus who perceives the image (i.e., to the existence of some sort of pernicious and question-begging "Cartesian Theatre"; see Dennett 1991). Rather, the idea is that images will be consumed and interpreted in whatever way perceptual states are consumed and interpreted.

Second, I don't, in fact, need to come down on one side or the other of the imagery debate. For my purposes, it matters little whether inner speech is relevantly like heard speech, or whether it is realized in complex mentalese *descriptions* of heard sentences instead. Either way, it can still be the case that the states in question bear a semantic interpretation. And either way, it can still be the case that certain kinds of cognitive process (namely, domain-integrative processes) are dependent on their existence.

### R5.3. Language and memory

**Schrauf** makes an interesting connection between the main thesis of the target article and evidence of the role of language in effortful cases of episodic memory building. For episodic memories will often need to link narrative information across different modular domains, and there is considerable evidence of the role of natural language in the process of reconstructing such memories.

Especially intriguing is the evidence that bilinguals find it easier to reconstruct a memory in the language that they would have been using at the time of the original events. I don't think **Schrauf**'s claim here is that memories are stored in one natural language rather than another; instead, it seems to be that there are *associations* of various sorts created between the currently used language and the various nonlinguistic components of episodic memory. This seems very plausible.

**Wynn & Coolidge** complain that it would have been helpful if my proposal had been compared with the model of working memory developed by Baddeley (1986),[14] and they then proceed to provide a very useful comparison of this sort for themselves. And I basically agree with their account of the connections between Baddeley's views and my own, including the role of the phonological loop, and their suggestion that Baddeley's central executive might be equated with the sort of "virtual" executive that gets appealed to on my own proposal.

I also think that **Wynn & Coolidge** are right to stress the importance of visual imagination (Baddeley's "visuo-spatial sketchpad") in the coordination of complex skilled action, as the work of Keller and Keller (1996) illustrates. I would suggest that visual imagination is one of the main vehicles of the sort of prelinguistic mental rehearsal discussed in section R1.2, indeed. But I would claim that prelinguistic visual imagery needs to be driven by the outputs of specific conceptual modules and/or by the contents of the practical reasoning module; and I would also claim that the process of generating images that link together the outputs of different conceptual modules is dependent on language.

### R5.4. Perspective taking and language

**MacWhinney** stresses the importance for central cognitive processes of perspective taking in production and comprehension of discourse. It is this, he says, which enables language to integrate information from many different parts of the brain. Taken in one way, this might seem to be a competitor to the hypothesis that it is sentences in LF that are the vehicles of intermodular integration. And this might indeed be what **MacWhinney** has in mind, because he (like **Baumeister & Vohs**) thinks that grammatical language is a late human *invention,* rather than an aspect of our innate human endowment (MacWhinney 2003). But if so, then the account founders on the evidence of an innately structured language faculty, as mentioned in section R4.6.

**MacWhinney**'s proposal might be better seen as an account of how *mental models* get built, I think; emphasizing their role in the comprehension process especially. If so, then it would emerge as a friendly supplement to my own account. For I, too, have stressed the importance of mental-model building in the comprehension of language. And it is this, in my view, that enables linguistic representations in LF to be taken as input and consumed by the various conceptual modules.

### R5.5. Vygotskian processing

Both **Frawley** and **Pléh** rightly point out that there are strong connections between my main thesis and the Vygotskian tradition in psychology and cognitive science. For this, too, stresses the importance of language (specifically, "silent speech") not just in development but also in mature cognition, especially when task demands increase. My intermodular integration proposal can be seen as one way of fleshing out what "hard" amounts to in this context.[15]

Both **Frawley** and **Pléh** point out, as well, that the use of inner speech in cognition isn't hardwired (so to speak), but is rather variable and opportunistic. With this, too, I agree. On my account, distinctively human creative thinking arises out of the interaction of a range of hardwired systems that were selected for other purposes. (These include a set of sophisticated conceptual modules for generating beliefs and desires; a language module that both takes inputs *from* the conceptual modules and provides inputs *to* them and that forms LF sentences integrating the outputs of the conceptual modules; feedback loops for mental rehearsal, which routinely run the contents of intended actions back though the conceptual modules for further evaluation; and a set of principles of relevance and coherence used in interpreting speech and evaluating the testimony of others.) What then had to be added were a set of *dispositions:* to generate novel linguistic contents as mere *suppositions,* for input to conceptual modules; to evaluate the consequences of those suppositions as if they were testimony; and to form explicit (language-involving) beliefs about appropriate patterns of reasoning and decision making.

I would predict that these dispositions, while to some degree universal amongst humans, are highly variable in their strength. Indeed, they might constitute the main components of the kind of "cognitive style" that Stanovich (1999) argues is the major determinant of differential success in batteries of reasoning tasks and measures of IQ.

### R5.6. Weak Whorfianism

In the target article I was concessive about the powers of different natural languages to sculpt cognition differently during development. And **Slobin** points out that, in the light of Levinson's results concerning the ways in which space is handled differently in different natural languages (Levinson 1996; 1997; 2002), I am therefore not entitled to assume that concepts such as *long wall on the left* will be cross-culturally available, existing prelinguistically and awaiting translation into language.[16]

Now, in one respect this is quite right. I very much doubt that the output of the geometry module will include the concepts of *left* and *right.* Acquiring the vocabulary of left and right would surely be a great deal easier for children if it did! So, in this respect the way in which I developed my hypothesis was misleading. Rather, the output of the geometry module might take the form, "The toy is at the corner of *this* [diagram] shape," which then has to be mapped laboriously into the language of left and right.

However, I now also believe that I may have been too concessive toward the "language sculpts cognition" approach (as **Atran**, too, points out). Each set of data will have to be examined on a case-by-case basis, of course. But in respect of Levinson's spatial results, Li and Gleitman (2002) provide an elegant demonstration that those results would seem to depend on the spatial cues that are salient in the testing environment. (However, see Levinson et al. 2002, for a vigorous reply.) Similarly, Papafragou et al. (2002) conduct a test of Slobin's (2002) claims concerning the ways in which different encodings of *manner of motion* in different languages will create patterns of attention that will lead, in turn, to effects on nonverbal memory. They are unable to find any such effects.

I think the jury is still out on the question whether language sculpts cognition, then. Certainly the weak forms in which this thesis is currently being pursued are consistent with the strong modularism adopted in my target article and also with my main thesis. But no such weak Whorfian claims are supported or entailed by my views. And the empirical data are still subject to a variety of interpretations.

Both **Frawley** and **Henser** claim that my view of the cognitive role of LF commits me to some sort of synchronic *linguistic relativity thesis.* They reason that if some thought is conducted in natural language, then (because languages differ in significant ways) people's thought processes, too, must differ in significant ways, depending on the language in which their thoughts are entertained. Such a claim is surely correct in respect of what Slobin (2002) calls "thinking for speaking." For if your language requires you to encode manner of motion, then of course, you have to think about manner in describing a motion event; and if your language uses cardinal directions to describe space, then of course, you will have to think in terms of such directions to tell someone where something is. But it is another matter to claim that speakers of different languages will differ in their *nonlinguistic* thoughts. And as previously indicated, I regard such claims as currently unproven.

Perhaps **Frawley** and **Henser** have in mind a stronger thesis, however. Perhaps their idea is that, by virtue of speaking different languages, certain thoughts will be inaccessible to one group of speakers that are available to the other. But it is highly controversial to claim that languages differ in their expressive powers. Indeed, the working hypothesis adopted by most linguists in the Chomskian tradition is that, at the deepest level, the LFs of any language are determined by UG (universal grammar) together with the relevant lexicon and parameter settings (Higginbotham 1995; Hornstein 1995). And it is generally assumed that there is no incommensurability between languages in respect of the thoughts that they can carry. (See Baker 2002 for extended and detailed discussion.)

### R5.7. Language and the transmission of culture

A number of commentators stress the importance of language in the generation and transmission of culture (**Baumeister & Vohs, Bryson, Hampton, Pléh**). I agree entirely. I didn't emphasize this in the target article, because my focus was on the synchronic cognitive functions of language. But nothing that I said was intended to be inconsistent with it.

Of course, language is the primary means by which we construct and communicate elements of culture, and it is also the primary means by which cultures are transmitted from one generation to the next, whether informally, through gossip and narrative, or formally, through education and training. And as **Baumeister & Vohs** point out, language is the vehicle for multigenerational accumulation and transmission of knowledge. Moreover, it is through language that most social processes are conducted and negotiated. And as **Hampton** highlights, too, language is crucial for the development of socially shared, explicit, logical, and scientific thinking. There is nothing here that I should want to disagree with.

## R.6. Conclusion

Modular models of mind are well motivated by the need to understand mental processes as computationally realized. But such models give rise to a problem: namely, to comprehend how flexible and creative human cognition could emerge out of the interactions of a set of modular components. The main thesis of the target article (that the language module serves to conjoin the contents of a suite of conceptual modules) is one aspect of a solution to that problem. And both that thesis and the evidence for it have, I believe, survived the commentary process clarified and further elaborated, but essentially undamaged.

ACKNOWLEDGMENT
I am grateful to the following for help and advice received while constructing this response to my commentators: Scott Atran, Stephen Crain, Paul Harris, Stephen Laurence, Gary Marcus, Paul Pietroski, Georges Rey, Michael Siegal, Barry Smith, and Helen Tager-Flusberg.

NOTES

**1.** In a way, Fodor doesn't deny either of the premises in this argument; nor does he deny that the conclusion follows. Rather, he believes that we have independent reasons to think that the conclusion is false; and he believes that we cannot even *begin* to see how amodular processes could be computationally realized. So he thinks that we had better give up attempting to do computational psychology (in respect of central cognition) for the foreseeable future. Fortunately, however, Fodor's reasons for thinking that central cognition is holistic in character are poor ones. For they depend on the assumption that scientific enquiry (social and public though it is) forms a useful model for the processes of individual cognition; and this supposition turns out to be incorrect (Carruthers 2003; but see also sect. R1.4 for discussion of a related problem for mental modularity).

**2.** A similar point holds in respect of the *outputs* of a module. Although we might expect that many modules will pass on their output to a quite restricted range of other systems, this is by no means mandatory. And the visual module, for example, would appear to make its outputs available to all central belief and desire-forming modules, as well as to the practical reasoning system to underpin indexical planning, in such a way as to ground thoughts of the form, "I'll go *that* way."

**3.** And incidentally, in claiming that *thoughts* are discrete structured states, I by no means intend to claim that all mental states share the same character. In particular, I believe that many *perceptual* states and processes are analog or continuous, rather than discrete (Carruthers 2000), just as **Dale & Spivey** assert. Of course there is *a* use of "thought" in which percepts can be described as "thoughts"; but this wasn't the usage I adopted in the target article. For me, thoughts are conceptualized propositional states, such as beliefs, judgments, and desires.

**4.** What *is* presupposed here, however, is that the representations in the practical reasoning system are in the right format for use in generating inputs to the various central-modular conceptual systems. This will be so if the practical reasoning system itself operates through the manipulation of sensory images; or if it uses conceptual structures which can exploit the same top-down systems deployed in the conceptualization of vision, in such a way as to generate visual images with appropriate contents, which can then be consumed by the central modules in the usual way. (I am assuming that nothing here yet involves the generation of cross-domain contents; that is supposed to depend on the activity of a language faculty and a language-dependent supposer [see sect. R1.4], to make their appearance rather later in evolution.) For such modules would, of course, have been set up already in such a way as to feed off perceptual outputs.

**5.** It is worth noting that a good deal of the evidence which has been cited (controversially) in support of chimpanzee theory of mind can also be used (much less controversially) to support the claim that this species of great ape, at least, engages in mental rehearsal. (I am quite prepared to believe that there is evidence of mental rehearsal from outside the great-ape lineage, too. But I shall not pursue the point here.) For example: A subordinate ape knows the location of some hidden food within an enclosure, and from previous experience expects to be followed by a dominant who will then take the food. So the subordinate heads off in the other direction and begins to dig. When the dominant pushes her aside and takes over the spot, she doubles back and retrieves and quickly eats the food. Such examples are generally discussed as providing evidence that chimps can engage in genuine (theory-of-mind involving) deception, that is, as showing that the chimp is intending to induce a false belief in the mind of another (Byrne 1995; Byrne & Whiten 1988), whereas critics have responded that chimpanzees may just be very smart behaviorists (Povinelli 2000; Smith 1996). But either way, it seems that the chimp must engage in mental rehearsal, predicting the effects of walking in the wrong direction and beginning to dig (the dominant will follow and take over the digging) and discerning the opportunities for hunger satisfaction that will then be afforded.

**6.** Note that on such an account the number module would be a bit like the cheater-detection system investigated by Cosmides & Tooby (1992). For the latter is supposed to operate on algorithms defined over *costs* and *benefits* in an exchange, where these can of course bridge other modular categories. A benefit can be a physical item (a jewel), a biological item (a goat), something socially defined (a wife), or something more psychological, like sympathy.

**7.** In a related error, **Molina** complains that language cannot be "the *locus* of thought" because thought organizes experience in terms of objects, events, properties, relations, and so on, not in terms of nouns, adjectives, and verbs. And he remarks (in intended criticism of me) that the mapping from conceptual category to syntactic category is many-to-many. I agree with these points entirely. The thesis of the target article isn't about concepts. I accept that most concepts are prior to and independent of language, indeed (in contrast with the view attributed to me by **Chater**). Rather, it concerns only a limited kind of thinking and reasoning (i.e., thinking that integrates the outputs of conceptual modules).

**8.** Note that a prediction of this account is that children will be capable of engaging in creative pretense only in respect of kinds of thing or activity for which they have names in their emerging vocabulary. I gather from Paul Harris (personal communication) that there are no data from normal children which address this question directly; and also from Michael Siegal (personal communication) that those working with late-signing deaf children haven't yet addressed this question, either.

**9.** **Samuels** also criticizes (at length) my claim that Hermer and Spelke's (1996) correlational data "strongly suggest" that language is implicated in cross-modular thought, on the grounds that correlations aren't evidence of causality. I agree with the latter point, of course; but then "suggests" neither means nor implies "is evidence of." And I believe that correlational data *do* strongly suggest causality, in the following sense: They render a causal hypothesis highly salient and a plausible candidate for further empirical testing.

**10.** I have been unable to find any data that speak to the question of whether subjects repair semantic errors while shadowing speech when undertaking a concomitant task. But it should also be noted that repairs aren't especially common, even when speech shadowing is the only task undertaken. Even the most common category of repair in Marslen-Wilson (1975) – namely, alterations of the third syllable of a semantically anomalous three-syllable word – was only found in 32 out of a total available 130 instances; and for most other categories of repair the proportions were *much* lower.

**11.** Note that **Gauker** *could* have defined conceptual thought in such a way as to warrant his conclusion. For example, on some construals of the *Generality Constraint* on concept-possession defended by Evans (1982), it might be required that thinkers should be capable of combining any concept together with any other concept into a single thought, to qualify as thinking conceptually. And then on my own account, too, nonlinguistic animals and prelinguistic children wouldn't count as engaging in conceptual thought. I just don't think that this would be a well-motivated definition of what it takes to think conceptually.

**12.** Recall that the vehicles of structured thoughts don't have to look much like sentences, on my view. So depending on the properties of the images in question and the manner of their manipulation, **Gauker's** idea might amount to the claim that animals (like humans, for much of the time) think through the exploitation of *mental models.* But a mental model can still be a conceptual thought.

**13.** In the case of **Clark** the connectionist framework is made explicit; in **O'Brien & Opie** it is largely hidden – but see O'Brien & Opie (2002) from which their comment is largely drawn.

**14.** Here I plead guilty. But let me say in mitigation that I did provide an extended comparison between Baddeley's views and the theory I was then developing (admittedly, not at that stage formulated in terms of modularity) in the final chapter of Carruthers (1996).

**15.** Of course, processing will also be hard which requires us to extend or correct the resources of just a single module, such as our folk-mechanics system. And processing will be *especially* hard where we are required, not just to *conjoin* contents from distinct modules, but to *integrate* those contents into a single system that may violate the assumptions made by one or another module; arguably this is the difficulty we face when we attempt to be physicalists about the mind. (This answers a question raised by **Hampton**.)

**16.** And note that if **Slobin** is right that there are languages that contain no equivalents of "left" and "right" (and no other means of linguistically representing egocentric spatial relations), then my prediction would be that such people will be incapable of solving the Spelke reorientation tasks. For under conditions of disorientation, absolute directions will no longer be available to them. So they will lack any way of combining landmark cues with the necessary spatial information in language.

## References

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


Allen, C. & Bekoff, M. (1997) *Species of mind: The philosophy and psychology of cognitive ethology.* MIT Press. [aPC]

Almor, A. (in press) Specialised behaviour without specialised modules. In: *Evolution and the psychology of thinking,* ed. D. E. Over. Psychology Press. [JSBTE]

Astington, J. (1996) What is theoretical about the child's theory of mind? In: *Theories of theories of mind,* ed. P. Carruthers & P. K. Smith. Cambridge University Press. [aPC]

Atran, S. (1990) *Cognitive foundations of natural history: Towards an anthropology of science.* Cambridge University Press. [aPC]

(1998) Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences* 21:547–68. [aPC]

(2002) An experimental approach to the cognitive basis of science: Universal and cultural factors in biological understanding. In: *The cognitive basis of science,* ed. P. Carruthers, S. Stich and M. Siegal. Cambridge University Press. [aPC]

(in press) Modest adaptationism (commentary). *Behavioral and Brain Sciences.* [SA]

Atran, S. & Lois, X. (2001) Reply to Shanker. *Current Anthropology* 42:498–500. [SA]

Baddeley, A. (1986) *Working memory.* Oxford University Press. [TW]

(1988) *Human memory.* Erlbaum. [aPC]

(1999) *Essentials of human memory.* Psychology Press. [WF]

(2001) Is working memory still working? *American Psychologist* 56:851–64. [TW]

Baddeley, A. & Logie, R. (1999) Working memory: The multiple-component model. In: *Models of working memory: Mechanisms of active maintenance and executive control,* ed. A. Miyake & P. Shah. Cambridge University Press. [TW, PC]

Baillargeon, R. (1995) Physical reasoning in infancy. In: *The cognitive neurosciences,* ed. M. Gazzaniga. MIT Press. [aPC]

Baker, M. (2002) *The atoms of language.* Basic Books. [rPC]

Balaban, M. & Waxman, S. (1997) Words may facilitate categorization in 9-month-old infants. *Journal of Experimental Child Psychology* 64:3–26. [FX]

Bannerman, D. M., Good, M. A., Butcher, S. P., Ramsay, M. & Morris, R. G. M. (1995) Distinct components of spatial learning revealed by prior training and NDMA receptor blockade. *Nature* 378:182–86. [JJB]

Barkow, J., Cosmides, L. & Tooby, J., eds. (1992) *The adapted mind.* MIT Press. [aPC]

Baron-Cohen, S. (1995) *Mindblindness.* MIT Press. [aPC]

Baron-Cohen, S. & Ring, H. (1994) A model of the mind-reading system: Neuropsychological and neuro-biological perspectives. In: *Children in Early Understanding of Mind* , ed. C. Lewis & P. Mitchell. Erlbaum. [rPC]

Barrouillet, P. & Lecas, J. F. (1998) How can mental models theory account for content effects in conditional reasoning? A developmental perspective. *Cognition* 67:209–53. [JSBTE]

Barsalou, L. W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22:577–609. [JAH]

Bates, E. (1999) Plasticity, localization, and language development. In: *The changing nervous system: Neurobehavioral consequences of early brain disorders,* ed. S. Broman & J. M. Fletcher. Oxford University Press. [JJB]

Baumeister, R. F. (in press) *The cultural animal: Human nature, meaning, and social life.* Oxford University Press. [RFB]

Bermudez, J. L. (2003) *Thinking without words.* Oxford University Press. [JLB]

Berry, D. C. & Dienes, Z. (1993) *Implicit learning.* Erlbaum. [JSBTE, PC]

Bickerton, D. (1981) *Roots of language.* Ann Arbor. [aPC, RV]

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

(1995) *Language and human behavior.* University of Washington Press. (UCL Press, 1996.) [DB, aPC]

(2002) From protolanguage to language. In: *The speciation of modern Homo sapiens,* ed. T. J. Crow. Oxford University Press. [DB]

Biegler, R. & Morris, R. G. (1993) Landmark stability is a prerequisite for spatial but not discrimination learning. *Nature* 361(6413):631–33. [LH-V]

(1996) Landmark stability: Studies exploring whether the perceived stability of the environment influences spatial representation. *Journal of Experimental Biology* 199(1):187–93. [LH-V]

Bloom, A. H. (1981) *The linguistic shaping of thought: A study in the impact of language on thinking in China and the West.* Erlbaum. [SH]

Bloom, P. (2000) *How children learn the meanings of words.* MIT Press. [rPC]

Bohnemeyer, J. (2003) The unique vector constraint: The impact of *direction* changes on the linguistic segmentation of motion events. In: *Representing direction in language and space,* ed. E. van der Zee & J. Slack. Oxford University Press. [DIS]

Bolinger, D. (1975) *Aspects of language,* 2nd edition. Harcourt Brace Jovanovich. [JAH]

Botterill, G. & Carruthers, P. (1999) *The philosophy of psychology.* Cambridge University Press. [aPC]

Bowerman, M. (1994) From universal to language-specific in early grammatical development. *Philosophical Transactions of the Royal Society, London, B* 346:37–45. [CP]

Bowerman, M. & Choi, S. (2001) Shaping meanings for language: Universal and language-specific in the acquisition and shaping of semantic categories. In: *Language acquisition and conceptual development,* ed. M. Bowerman & S. Levinson. Cambridge University Press. [AC, DIS]

Bowerman, M. & Levinson, S., eds. (2001) *Language acquisition and conceptual development.* Cambridge University Press. [aPC]

Boysen, S. T., Bernston, G., Hannan, M. & Cacioppo, J. (1996) Quantity-based inference and symbolic representation in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes* 22:76–86. [JJB]

Brown, R. W. & Lenneberg, E. H. (1954) A study in language and cognition. *Journal of Abnormal and Social Psychology* 49:454–62. [CP]

Bruner, J. S., Olver, R. R. & Greenfield, P. M. (1966) *Studies in cognitive growth.* Wiley. [JAH]

Bryson, J. (2000) Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental and Theoretical Artificial Intelligence* 12:165–90. [rPC]

Bryson, J. J. (2001). *Intelligence by design: Principles of modularity and coordination for engineering complex adaptive agents.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA. [JJB]

Butterworth, B. (1999) *What counts: How every brain is hardwired for math.* Simon and Schuster. [rPC]

Byrne, R. (1995) *The thinking ape.* Oxford University Press. [aPC]

Byrne, R. & Whiten, A., eds. (1988) *Machiavellian intelligence.* Oxford University Press. [aPC]

(1998) *Machiavellian intelligence II: Evaluations and extensions.* Cambridge University Press. [aPC]

Calvin, W. H. & Bickerton, D. (2000) *Lingua ex machina.* MIT Press. [DB]

Canedo, A. (1997) Primary motor cortex influences on the descending and ascending systems. *Progress in Neurobiology* 51(3):287–335. [LH-V]

Carey, S. (1985) *Conceptual change in childhood.* MIT Press. [aPC]

(2001) Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind and Language* 16:37–55. [rPC]

Carey, S. & Spelke, E. (1994) Domain-specific knowledge and conceptual change. In: *Mapping the mind,* ed. L. Hirshfeld & S. Gelman. Cambridge University Press. [aPC]

Carroll, J. B. & Casagrande, J. B. (1958) The function of language classification in behavior. In: *Readings in social psychology,* ed. E. Maccoby, F. Newcomb & E. Hartley. Holt. [CP]

Carruthers, P. (1992) *Human knowledge and human nature.* Oxford University Press. [aPC]

(1996) *Language, thought and consciousness.* Cambridge University Press. [aPC, SH, PPS]

(1998a) Thinking in language? Evolution and a modularist possibility. In: *Language and thought,* ed. P. Carruthers & J. Boucher. Cambridge University Press. [aPC, SH]

(1998b) Conscious thinking: Language or elimination? *Mind and Language* 13:323–42. [aPC]

(2000) *Phenomenal consciousness: A naturalistic theory.* Cambridge University Press. [aPC, KF]

(2002a) The roots of scientific reasoning: Infancy, modularity and the art of tracking. In: *The Cognitive basis of science,* ed. P. Carruthers, S. Stich & M. Siegal. Cambridge University Press. [aPC]

(2002b) Human creativity: Its evolution, its cognitive basis, and its connections with childhood pretence. *British Journal for the Philosophy of Science* 53:1–25. [rPC]

(2003a) Is the mind a system of modules shaped by natural selection? In: Contemporary debates in the philosophy of science, ed. C. Hitchcock. Blackwell. [Retrieved November 14, 2002, from *http://www.philosophy.umd.edu/people/faculty/pcarruthers/Shaped-modules.htm*] [SH]

(2003b) Moderately massive modularity. In: *Mind and persons,* ed. A. O'Hear. Cambridge University Press. [rPC]

(in press a) Distinctively human thinking: Modular precursors and components. In: *The structure of the innate mind,* ed. P. Carruthers, S. Laurence & S. Stich. [Retrieved November 16, 2002, from *http://www.philosophy.umd.edu/people/faculty/pcarruthers/Distinctively-human.htm*] [SH]

(in press b) On Fodor's problem. *Mind and Language.* [rPC, SH]

(in press c) Practical reasoning in a modular mind. [rPC]

Chater, N. (1999) The search for simplicity: A fundamental cognitive principle? *Quarterly Journal of Experimental Psychology,* 52A:273–302. [aPC]

Cheng, K. (1986) A purely geometric module in the rat's spatial representation. *Cognition* 23:149–78. [aPC, LVH]

Choi, S. & Bowerman, M. (1991) Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41:83–121. [DIS]

Choi, S. & Gopnik, A. (1995) Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language* 22:497–529. [aPC]

Chomsky, N. (1976) *Reflections on language.* Temple Smith. [aPC]

(1988) *Language and problems of knowledge.* MIT Press. [aPC]

(1995) *The minimalist program.* MIT Press. [aPC, NC, SH, CM]

(2000) *New horizons in the study of language and mind.* Cambridge University Press. [rPC]

Churchland, P. M. (2001) Neurosemantics: On the mapping of minds and the portrayal of worlds. [GO]

Clark, A. (1998) Magic words: How language augments human computation. In: *Language and thought,* ed. P. Carruthers & J. Boucher. Cambridge University Press. [aPC]

Clements, W. A. & Perner, J. (1994) Implicit understanding of belief. *Cognitive Development* 9:377–95. [JGdV]

Cohen, L. J. (1993) *An essay on belief and acceptance.* Oxford University Press. [aPC]

Coltheart, M. (1999) Modularity and cognition. *Trends in Cognitive Sciences* 3:115–20. [RV]

Conway, M. A. (1996) Autobiographical knowledge and autobiographical memories. In: *Remembering our past: Studies in autobiographical memory,* ed. D. C. Rubin. Cambridge University Press. [RWS]

(2002) Sensory-perceptual episodic memory and its context: Autobiographical memory. In: *Episodic memory: New directions in research,* ed. A. Baddeley, J. P. Aggleton & M. A. Conway. Oxford University Press. [RWS]

Cosmides, L. & Tooby, J. (1992) Cognitive adaptations for social exchange. In: *The adapted mind,* ed. J. Barkow, L. Cosmides & J. Tooby. Oxford University Press. [aPC]

Cummins, R. (1996) *Representations, targets and attitudes*. MIT Press. [GO]

Curtiss, S. (1977) *Genie: A psycholinguistic study of a modern-day wild child.* Academic Press. [aPC]

Davidson, D. (1975) Thought and talk. In: *Mind and language,* ed. S. Guttenplan. Oxford University Press. [aPC]

(1982) Rational animals. In: *Actions and events,* ed. E. Lepore & B. McLaughlin. Blackwell. [aPC]

Dawkins, R. (1976) *The selfish gene.* Oxford University Press. [aPC]

De Villiers, J. (2000) Language and theory of mind: What is the developmental relationship? In: *Understanding other minds: Perspectives from autism and developmental cognitive neuroscience,* ed. S. Baron-Cohen, H. Tager-Flusberg & D. Cohen. Cambridge University Press. [aPC]

De Villiers, J. G. & de Villiers, P. A. (2000) Linguistic determinism and false belief. In: *Children's reasoning and the mind,* ed. P. Mitchell & K. Riggs. Psychology Press. [JGdV]

(2002) Language for thought: Coming to understand false beliefs. In: *Whither Whorf?,* ed. D. Gentner & S. Goldin-Meadow. MIT Press. [JGdV]

De Villiers, J. G. & Pyers, J. (1997) Complementing cognition: The relationship between language and theory of mind. In: *Proceedings of the 21st Annual Boston University Conference on Language Development.* Cascadilla Press. [JGdV]

(2002) Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief understanding. *Cognitive Development* 17:1037–60. [JGdV]

De Waal, F. (1982) *Chimpanzee politics.* Jonathan Cape. [aPC]

(1996) *Good natured.* Harvard University Press. [aPC]

DeWaal, F. B. M. (2001) *The ape and the sushi master: Reflections of a primatologist.* Basic Books. [RFB]

Dehaene, S. (1997) *The number sense.* Oxford University Press. [aPC]

Dennett, D. (1983) Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences* 6:343–55. [JGdV]

(1991) *Consciousness explained.* Penguin Press. [aPC, GO]

(1991) *Consciousness explained.* Little, Brown. [JJB, KF, PPS]

Diaz R. & Berk, L., eds. (1992) *Private speech: From social interaction to self-regulation.* Erlbaum. [aPC]

Dickinson, A. (1994) Instrumental conditioning. In: *Animal learning and cognition,* ed. N. Mackintosh. Academic Press. [aPC]

Dickinson, A. & Balleine, B. (2000) Causal cognition and goal-directed action. In: *The evolution of cognition,* ed. C. Heyes & L. Huber. MIT Press. [aPC]

Dickinson, A. & Shanks, D. (1995) Instrumental action and causal representation. In: *Causal cognition,* ed. D. Sperber, D. Premack & A. Premack. Blackwell. [aPC, rPC]

Dienes, Z. & Perner, J. (1999) A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences* 22:735–808. [JGdV]

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]

Donald, M. (1991) *Origins of the modern mind.* Harvard University Press. [KN, CP]

(1993) Précis of *Origins of the modern mind*: Three stages in the evolution of culture and cognition. *Behavioral and Brain Sciences* 16:737–91. [CP]

(1998) Mimesis and the executive suite: Missing links in language evolution. In: *Approaches to the evolution of language,* ed. J. R. Hurford, M. G. Studdert-Kennedy & C. Knight. Cambridge University Press. [BM]

(2001) *A mind so rare: The evolution of human consciousness.* Norton. [CP]

Dougherty, J. & Keller, C. (1982) Taskonomy: A practical approach to knowledge structures. *American Ethnologist* 9:763–74. [TW]

Dowty, A. D. R., Wall, A. R. E. & Peters, P. S. (1981) *Introduction to Montague Semantics.* Reidel. [NC]

Dummett, M. (1981) *The interpretation of Frege's philosophy.* Duckworth. [aPC]
(1989) Language and communication. In: *Reflections on Chomsky,* ed. A. George. Blackwell. [aPC]
(1994) *Origins of analytical philosophy.* Harvard University Press. [aPC]
Edelman, S. (1998) Representation is representation of similarities. *Behavioral and Brain Sciences* 21:449–98. [PPS]
Emmorey, K. (1998) Mental rotation within linguistic and nonlinguistic domains in users of American Sign Language. *Cognition* 68(3):221–46. [SH]
Ervin-Tripp, S. M. (1961) Learning and recall in bilinguals. *American Journal of Psychology* 74:446–51. [SH]
(1964a) An analysis of the interaction of language, topic and listener. *American Anthropologist* 66(6): 86–102 (special publication: *The Ethnography of Communication,* ed. J. J. Gumperz & D. Hymes). [SH]
(1964b) Language and TAT content in bilinguals. *Journal of Abnormal and Social* Psychology 68:500–507. [SH]
Evans, G. (1982) *The varieties of reference.* Oxford University Press. [rPC]
Evans, J. & Over, D. (1996) *Rationality and reasoning.* Psychology Press. [aPC]
Evans, J. St. B. T., Handley, S. H. & Over, D. E. (in press) Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition.* [JSBTE]
Evans, J. St. B. T., Newstead, S. E. & Byrne, R. M. J. (1993) *Human reasoning: The psychology of deduction.* Erlbaum. [JSBTE]
Evans, J. St. B. T. & Over, D. E. (1996) *Rationality and reasoning.* Psychology Press. [JSBTE]
Evans, J. St. B. T., Over, D. E. & Handley, S. H. (in press) A theory of hypothetical thinking. In: *Decision making and reasoning,* ed. D. Hardman & L. Maachi. Wiley. [JSBTE]
Fauconnier, G. & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities.* Basic Books. [GAB]
Ferreira, F., Bailey, K. G. D. & Ferraro, V. (2002) Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11:11–15. [RD]
Fiddick, L., Cosmides, L. & Tooby, J. (2000) No interpretation without representation: The role of domain-specific representations and inferences in the Wason selection task. *Cognition* 77:1–79. [aPC]
Fodor, J. (1975) *The language of thought.* Harvester. [rPC, NC, JGdV, CP]
(1980) On the impossibility of acquiring "more powerful" structures. In: *Language and learning: The debate between Jean Piaget and Noam Chomsky,* ed. M. Piattelli-Palmarini. Harvard University Press. [NC]
(1981) The present status of the innateness controversy. In: *Representations.* Harvester Press. [aPC]
(1983) *The modularity of mind.* MIT Press. [aPC, CP, RS, RV]
(1998) *Concepts: Where cognitive science went wrong.* Clarendon Press. [JAH]
(2000) *The mind doesn't work that way: The scope and limits of computational psychology.* MIT Press. [rPC, RD, PR]
Fodor, J. & McLaughlin, B. (1990) Connectionism and the problem of systematicity. *Cognition* 35:183–204. [rPC]
Fodor, J. & Pylyshyn, Z. (1988) Connectionism and cognitive architecture. *Cognition* 28:3–71. [rPC, NC]
Fong, G. T., Krantz, D. H. & Nisbett, R. E. (1986) The effects of statistical training on thinking about everyday problems. *Cognitive Psychology* 18:253–92. [JSBTE]
Frankish, K. (1998a) Natural language and virtual belief. In: *Language and thought,* ed. P. Carruthers & J. Boucher. Cambridge University Press. [aPC, KF, TW]
(1998b) A matter of opinion. *Philosophical Psychology* 11:423–42. [aPC, KF, TW]
(in press) *Mind and supermind.* Manuscript. [aPC, KF]
Frawley, W. (1997) *Vygotsky and cognitive science: Language and the unification of the social and computational mind.* Harvard University Press. [WF]
(1999) Vygotsky and developmental computationalism: Arguments for a language *for* thought. In: *Language and thought in development: Cross-linguistic studies,* ed. P. Broeder & J. Muure. Gunter Narr. [WF]
(2002a) Control and cross-domain mental computation: Evidence from language breakdown. *Computational Intelligence* 18:1–28. [WF]
(2002b) Mental computation and language breakdown: Clarifications, extensions, and responses: Evidence from language breakdown. *Computational Intelligence* 18:59–86. [WF]
Frawley, W. & Smith, R. N. (2001) A processing theory of alexithymia. *Journal of Cognitive Systems Research* 2:189–206. [WF]
Frege, G. (1918/1979) The thought. In: *Posthumous writings,* trans. P. Long & R. White. Blackwell. [rPC]
Gallistel, R. (1990) *The organization of learning.* MIT Press. [aPC, LH-V]
(2000) The replacement of general-purpose learning models with adaptively specialized learning modules. In: *The new cognitive neurosciences,* 2nd Edition. ed. M. Gazzaniga. MIT Press. [rPC]
Gallistel, R. & Gellman, R. (1992) Preverbal and verbal counting and computation. *Cognition* 44:43–74. [aPC]
Gallistel, R. & Gibson, J. (2001) Time, rate and conditioning. *Psychological Review* 108:289–344. [rPC]
Gauker, C. (2003) *Words without meaning.* MIT Press. [CG]
Gazzaniga, M. (1998) *The mind's past.* California University Press. [rPC]
Gelman, S. (2003) *The essential child: Origins of essentialism in everyday thought.* Oxford University Press. [FX]
Gentner, D. & Bowles, B. (2001) Convention, form, and the processing of figurative language. *Metaphor and Symbol* 16:223–47. [GAB]
Gentner, D., Holyoak, K. J. & Kokinov, B., eds. (2001) *The analogical mind: Perspectives from cognitive science.* MIT Press. [GAB]
Gernsbacher, M. A. (1990) *Language comprehension as structure building.* Erlbaum. [BM]
Gibbs, R. (1994) *The poetics of mind: Figurative thought, language, and understanding.* Cambridge University Press. [GAB]
Gibbs, R., Lima, P. & Francuzo, E. (in press) Metaphor is grounded in embodied experience. *Journal of Pragmatics.* [GAB]
Gibson, B. M., Shettleworth, S. J. & McDonald, R. J. (2002) Finding a goal on dry land and in the water: Differential effects of disorientation on spatial learning. *Behavioral Brain Research* 123(1):103–11. [LH-V]
Gick, M. L. & Holyoak, K. J. (1983) Schema induction and analogical transfer. *Cognitive Psychology,* 15:1–38. [PFD]
Gigerenzer, G. & Todd, P. (2000) *Simple heuristics that make us smart.* Oxford University Press. [rPC]
Goldman, A. (1989) Interpretation psychologized. *Mind and Language* 4:161–85. [aPC]
(1993) The psychology of folk-psychology. *Behavioral and Brain Sciences* 16:15–28. [aPC]
Gómez, J. (1998) Some thoughts about the evolution of LADS, with special reference to TOM and SAM. In: *Language and thought,* ed. P. Carruthers & J. Boucher. Cambridge University Press. [aPC]
Gopnik, A. (1993) How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences* 16:1–14. [rPC]
(2001) Theories, language and culture. In: *Language acquistion and conceptual development,* ed. M. Bowerman & S. Levinson. Cambridge University Press. [aPC]
Gopnik, A., Choi, S. & Baumberger, T. (1996) Cross-linguistic differences in early semantic and cognitive development. *Cognitive Development* 11:197–227. [aPC]
Gopnik, A. & Melzoff, A. (1997) *Words, thoughts and theories.* MIT Press. [aPC]
Gordon, R. (1986) Folk psychology as simulation. *Mind and Language* 1:158–71. [aPC]
(1995) Simulation without introspection or inference from me to you. In: *Mental simulation: Evaluations and applications,* ed. T. Stone & M. Davies. Blackwell. [aPC]
Gouteux, S., Thinus-Blanc, C. & Vauclair, S. (in press) Rhesus monkeys use geometric and non-geometric information during a reorientation task. *Journal of Experimental Psychology: Gen. Proc.* 130:505–19. [aPC, LVH, GV]
Grady, J. (1999) A typology of motivation for conceptual metaphor: Correlation vs. resemblance. In: *Metaphor in cognitive linguistics,* ed. R. Gibbs & G. Steen. Benjamins. [GAB]
Greenberg, D. & Rubin, D. C. (in press) The neuropsychology of autobiographical memory. *Cortex.* [RWS]
Grosjean, F. (1982) *Life with two languages.* Harvard University Press. [SH]
Gurney, K., Prescott, T. J. & Redgrave, P. (1998) The basal ganglia viewed as an action selection device. In: *Proceedings of the International Conference on Artificial Neural Networks, Skövde, Sweden.* [JJB]
Hale, C. & Tager-Flusberg, H. (2003) The influence of language on theory of mind. *Developmental Science* 7. [rPC]
Harris, P. (1996) Desires, beliefs and language. In: *Theories of Theories of Mind,* ed. P. Carruthers & P. K. Smith. Cambridge University Press. [aPC]
(2000) *The work of the imagination.* Blackwell. [aPC]
(2002) What do children learn from testimony? In: *The cognitive basis of science,* ed. P. Carruthers, S. Stich & M. Siegal. Cambridge University Press. [aPC]
Hauser, M. (2000) *Wild minds.* Penguin Press. [aPC]
Hauser, M. & Carey, S. (1998) Building a cognitive creature from a set of primitives. In: *The evolution of mind,* ed. D. Cummins & C. Allen. Oxford University Press. [aPC]
Hauser, M. D. (1999) Perseveration, inhibition and the prefrontal cortex: A new look. *Current Opinion in Neurobiology* 9:214–22. [JJB]
Heal, J. (1986) Replication and functionalism. In: *Language, mind and logic,* ed. J. Butterfield. Cambridge University Press. [aPC]
(1995) How to think about thinking. In: *Mental simulation: Evaluations and applications,* ed. T. Stone & M. Davies. Blackwell. [aPC]
Hemphill, R. (1971) Auditory hallucinations in polyglots. *South African Medical Journal* 18:1391–94. [SH]

Henser, S. J. (2002a) Thinking in Japanese? What have we learned about language-specific thought since Ervin-Tripp's 1964 psychological tests of Japanese-English bilinguals? In: *Nissan Occasional Paper Series* 32. Nissan Institute of Japanese Studies, Oxford University. [Full text available as a PDF file at http://www.nissan.ox.ac.uk/nops/nops32.pdf] [SH]
(2000b) Natural language use in habitual propositional-type thought: Support from Japanese-English and English-Japanese bilingual covert codeswitching data. Unpublished doctoral dissertation, University of London. [British Library shelf mark DX217428.] [SH]
Hermer, L. (1993) Increases in cognitive flexibility over development and evolution: Candidate mechanisms. *Proceedings of the Cognitive Science Society.* [LH-V]
Hermer, L. & Spelke, E. (1994) A geometric process for spatial reorientation in young children. *Nature* 370:57–59. [aPC, LH-V, GV]
(1996) Modularity and development: The case of spatial reorientation. *Cognition* 61:195–232. [aPC, LH-V, PR, RS, GV]
Hermer-Vazquez, L., Moffet, A. & Munkholm, P. (2001) Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition* 79(3):263–99. [LH-V, RS]
Hermer-Vazquez, L., Spelke, E. & Katsnelson, A. (1999) Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology* 39:3–36. [JJB, JLB, aPC, GO, PR, RS, RWS, GV, RV, TW]
Herrnstein, R. J. (1992) Levels of stimulus-control: A functional approach. In: *Animal Cognition.,* ed. C. R. Gallistel. MIT Press, 133–66. [CG]
Heyes, C. M. (1998) Theory of mind in nonhuman primates. *Behavioral and Brain Sciences* 21:101–14. [CG]
Higginbotham, J. (1995) *Sense and syntax.* Oxford University Press. (An inaugural lecture delivered before the University of Oxford on October 20, 1994.) [rPC]
Hirschfeld, L. & Gelman, S., eds. (1994) *Mapping the mind: Domain specificity in cognition and culture.* Cambridge University Press. [aPC]
Hoen, M. & Dominey, P. F. (2000) ERP analysis of cognitive sequencing: A left anterior negativity related to structural transformation processing. *NeuroReport* 11(14):3187–91. [PFD]
Hoen, M., Golembiowski, M., Guyot, E., Deprez, V., Caplan, D. & Dominey, P. F. (2003) Training with cognitive sequences improves syntactic comprehension in agrammatic aphasics. *NeuroReport* [PFD]
Horgan T. & Tienson, J. (1996) *Connectionism and philosophy of psychology.* MIT Press. [aPC, CG]
Hornstein, N. (1995) *Logical form.* Blackwell. [rPC]
Hughes, C. & Plomin, R. (2000) Individual differences in early understanding of mind: Genes, nonshared environment and modularity. In: *Evolution and the human mind,* ed. P. Carruthers & A. Chamberlain. Cambridge University Press. [aPC]
Hughes, L. (1993) *ChimpWorld: A wind-tunnel for the social sciences.* Ph.D. dissertation, Yale University. [aPC]
Hunt, E. & Agnoli, F. (1991) The Whorfian hypothesis, a cognitive psychology perspective. *Psychological Review* 98(3):377–89. [SH]
Hurford, J. R. (2002) The neural basis of predicate-argument structure. *Behavioral and Brain Sciences* (http://www.bbsonline.org/Preprints/Hurford/). [JRH]
Hurlburt, R. (1990) *Sampling normal and schizophrenic inner experience.* Plenum Press. [aPC]
(1993) *Sampling inner experience with disturbed affect.* Plenum Press. [aPC]
Jackendoff, R. (1983) *Semantics and cognition.* MIT Press. [CM]
(1992) *Languages of the mind: Essays on mental representation.* MIT Press. [CP]
(1996) How language helps us think. *Pragmatics and Cognition* 4(1):1–34. [JGdV]
(1997) *The architecture of the language faculty.* MIT Press. [CM, PPS]
(2002) *Foundations of language: Brain, meaning, grammar and evolution.* Oxford University Press. [SH, CM]
Jellinek, A. (1977) The lower Paleolithic. *Annual Review of Anthropology,* 6:1–32. [DB]
Jepson, D., Krantz, D. H. & Nisbett, R. E. (1983) Inductive reasoning: Competence or skill? *Behavioral and Brain Sciences* 6:494–501. [JSBTE]
Johnson-Laird, P. (1983) *Mental models.* Cambridge University Press. [aPC, JSBTE]
(1996) Space to think. In: *Language and Space,* ed. P. Bloom, M. A. Peterson, L. Nadel & M. F. Garret. MIT Press. [PFD]
Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hove, UK: Lawrence Erlbaum Associates.
Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics and inference. *Psychological Review,* 109, 646–78.
Kaas, J. H. (1999) The transformation of association cortex into sensory cortex. *Brain Research Bulletin* 50:425. [RV]
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive change.* MIT Press. [JJB, rPC, KN, CP]
Katz, A., Cacciari, C., Gibbs, R. & Turner, M. (1997) *Figurative language and thought.* Oxford University Press. [GAB]
Keller, C. & Keller, J. (1996) *Cognition and tool use: The blacksmith at work.* Cambridge University Press. [TW]
Kelly, D. M., Spetch, M. L. & Heth, C. D. (1998) Pigeons (*Columba livia*) encoding of geometric and featural properties of a spatial environment. *Journal of Comparative Psychology* 112:259–69. [GV]
Kertesz, A. (1988) Cognitive function in severe aphasia. In: *Thought without language,* ed. L. Weiskrantz. Oxford University Press. [aPC]
Kingsbury, M. A. & Finlay, B. L. (2001) The cortex in multidimensional space: Where do cortical areas come from? *Development Science* 4:125–57. [RD]
Koslowski, B. (1996) *Theory and evidence.* MIT Press. [aPC]
Kosslyn, S. (1994) *Image and brain.* MIT Press. [aPC, TW]
Kosslyn, S. M. (1980) *Image and mind.* Harvard University Press. [PPS]
Kroll, J. F. & de Groot, A. M. B. (1997) Lexical and conceptual memory in the bilingual: Mapping form to memory in two languages. In: *Tutorials in bilingualism: Psycholinguistic perspectives,* ed. A. M. B. de Groot & J. F. Kroll. Erlbaum. [RWS]
Kuhlmeier, V. A. & Boysen, S. T. (2002) Chimpanzees (*Pan troglodytes*) recognize spatial and object correspondences between a scale model and its referent. *Psychological Science* 13(1):60–63. [JJB]
Kuperstein, M. (1988) Neural model of adaptive hand-eye coordination for single postures. *Science* 239(4845):1308–11. [PFD]
Lakatos, I. (1970) The methodology of scientific research programmes. In: *Criticism and the growth of knowledge,* ed. I. Lakatos & A. Musgrave. Cambridge University Press. [aPC]
Lakoff, G. (1987) *Women, fire and dangerous things: What categories reveal about the mind.* University of Chicago Press. [SH]
Lakoff, G. & Johnson, M. (1980) *Metaphors we live by.* Chicago University Press. [rPC]
(1999) *Philosophy in the flesh.* Basic Books. [GAB, PC]
Larson, R. & Segal, G. (1995) *Knowledge of language: An introduction to semantic theory.* MIT Press. [PR]
Larsen, S., Schrauf, R. W., Fromholt, P. & Rubin, D. C. (2002) Inner speech and bilingual autobiographical memory: A Polish-Danish cross-cultural study. *Memory* 10(1):45–54. [RWS]
Laski, E. & Taleporos, E. (1977) Anticholinergic psychosis in a bilingual: A case study. *American Journal of Psychiatry* 134:1038–40. [SH]
Laurence, S. & Margolis, E. (2001) The poverty of the stimulus argument. *British Journal for the Philosophy of Science* 52:217–76. [rPC]
Learmonth, A., Nadel, L. & Newcombe, N. (in press) Children's use of landmarks: Implications for modularity theory. *Psychological Science.* [aPC, GV]
Learmonth, A., Newcombe N. & Huttenlocher J. (2001) Toddlers' use of metric information and landmarks to reorient. *Journal of Experimental Child Psychology,* 80:225–44. [aPC, LH-V]
Leslie, A. (1994) ToMM, ToBY and Agency: Core architecture and domain specificity. In: *Mapping the mind,* ed. L. Hirschfeld & S. Gelman. Cambridge University Press. [aPC]
Levelt, W. (1989) *Speaking: From intention to articulation.* MIT Press. [aPC, KF, CM, RS, PC]
Levinson, S. C. (1996) Frames of reference and Molyneux's question: Cross-linguistic evidence. In: *Language and space,* ed. P. Bloom, M. Peterson, L. Nadel & M. Garrett. MIT Press. [DIS, PC]
(1997) From outer space to inner space: Linguistic categories and nonlinguistic thinking. In: *Language and conceptualization,* ed. E. Pederson & J. Nuyts. Cambridge University Press. [SH, DIS, PC]
(2002) *Space in language and cognition: Explorations in linguistic diversity.* Cambridge University Press. [DIS, PC]
Levinson, S., Kita, S., Haun, D. & Rasch, B. (2002) Returning the tables: Language effects and spatial reasoning. *Cognition* 84:155–88. [rPC]
Li, P. & Gleitman, L. (2002) Turning the tables: Language and spatial reasoning. *Cognition* 83:265–94. [rPC]
Li, P., Farkas, I. & MacWhinney, B. (in press) The origin of categorical representations of language in the brain. *Cognitive Science.* [BM]
Liebenberg, L. (1990) *The art of tracking: The origin of science.* David Philip. [aPC]
Livesey, P. J. (1986) *Learning and emotion: A biological synthesis, vol. 1, Evolutionary processes.* Erlbaum. [JJB]
Lobe, C. G. & Nagy, A. 1998. Conditional genome alteration in mice. *Bioessays* 20(3):200–8. [LH-V]
Logie, R. (1995) *Visuo-spatial working memory.* Erlbaum. [TW]
Lucy, J. (1992a) *Language, diversity and thought: A reformulation of the linguistic relativity hypothesis.* Cambridge University Press. [aPC]
(1992b) *Grammatical categories and cognition: A case-study of the linguistic relativity hypothesis.* Cambridge University Press. [aPC]
Lucy, J. & Gaskins, S. (2001) Grammatical categories and development of classification preferences: A comparative approach. In: *Language acquisition and conceptual development,* ed. M. Bowerman & S. Levinson. Cambridge University Press. [aPC, AC]

Maaswinkel, H. & Whishaw, I. Q. (1999) Homing with locale, taxon, and dead reckoning strategies by foraging rats: Sensory hierarchy in spatial navigation. *Behavioral Brain Research* 99(2):143–52. [LH-V]
MacWhinney, B. (1977) Starting points. Language, 53:152–68. [BM]
(1999) The emergence of language from embodiment. In: *The emergence of language,* ed. B. MacWhinney. Erlbaum. [BM]
(2003) The gradual evolution of language. In: *The evolution of language,* ed. B. Malle & T. Givón. Benjamins. [BM]
(in press) Perspective taking and grammar. *Behavioral and Brain Sciences.* [BM]
Magee, B. (1997) *Confessions of a philosopher.* Weidenfeld & Nicolson. [PPS]
Malson, L. (1972) *Wolf children and the problem of human nature.* Monthly Review Press. [aPC]
Marcus, G. (2001) *The algebraic mind.* MIT Press. [rPC]
Marian, V. & Neisser, U. (2000) Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General* 129(3):361–68. [RWS]
Marslen-Wilson, W. (1975) Sentence perception as an interactive parallel process. *Science* 189:226–28. [rPC, RD]
Maynard Smith, J. & Szathmary, E. (1995) *The major transitions in evolution.* W. H. Freeman. [DB]
McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. (1995) Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* 102(3):419–57. [JJB]
McDermott, D. (2001) *Mind and mechanism.* MIT Press. [rPC]
McDonald, R. J. & White, N. M. (1993) A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience* 107(1): 3–22. [LH-V]
McDowell, J. (1994) *Mind and world.* MIT Press. [aPC]
McGuigan, F. J. (1978) *Cognitive Psychophysiology: Principles of Covert Behavior.* Prentice Hall. [CP]
Medin, D. & Ortony, A. (1989) Psychological essentialism. In: *Similarity and analogical reasoning,* ed. S. Vosniadou & A. Ortony. Cambridge University Press. [FX]
Minsky, M. (1985) *The society of mind.* Simon and Schuster. [JJB]
Mithen, S. (1990) *Thoughtful foragers: A study of prehistoric decision making.* Cambridge University Press. [aPC]
(1996) *The prehistory of the mind.* Thames Hudson. [JLB, aPC, JRH]
Miyake, A. & Shah, P. (1999) *Models of working memory: Mechanisms of active maintenance and executive control.* Cambridge University Press. [TW]
Moscovitch, M. (1992) Memory and working-with memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience* 4:257–67. [RWS]
Nelson, K. (1986). *Event knowledge: Structure and function in development.* Erlbaum. [KN]
(1996) *Language in cognitive development: The emergence of the mediated mind.* Cambridge University Press. [aPC, KN]
(in press) Language pathways to the community of minds. In: *Why language matters to theory of mind,* ed. J. W. Astington & J. Baird. Cambridge University Press. [KN]
Nelson, K., Hampson, J. & Kessler Shaw, L. (1993) Nouns in early lexicons: Evidence, explanations, and implications. *Journal of Child Language* 20:61–84. [KN]
Nelson, K., Plesa, D., Goldman, S., Henseler, S., Presler, N. & Walkenfeld, F. F. (2003) Entering a community of minds: An experiential approach to theory of minds. *Human Development,* 191:1–23. [KN]
Nichols, S. & Stich, S. (1998) Theory-theory to the max. *Mind and Language* 13:421–49. [aPC]
(2000) A cognitive theory of pretense. *Cognition* 74:115–47. [rPC]
(forthcoming) *Mindreading.* Oxford University Press. [aPC]
Nisbett, R. & Ross, L. (1980) *Human inference.* Prentice Hall. [rPC]
Nisbett, R. & Wilson, T. (1977) Telling more than we can know. *Psychological Review* 84:231–95. [rPC]
Oaksford, M. & Chater, N. (1991) Against logicist cognitive science. *Mind and Language* 6:1–38. [NC]
(1998) *Rationality in an uncertain world: Essays on the cognitive science of human reasoning.* Psychology Press. [NC]
O'Brien, G. & Opie, J. (2002) Radical connectionism: Thinking with (not in) language. *Language and Communication* 22:313–29. [GO, PC]
Olson, D. R. (1994) *The world on paper: The conceptual and cognitive implications of writing and reading.* Cambridge University Press. [JGdV]
Origgi, G. & Sperber, D. (2000) Evolution, communication and the proper function of language. In: *Evolution and the Human Mind,* ed. P. Carruthers & A. Chamberlain. Cambridge University Press. [rPC]
Over, D. E. & Evans, J. St. B. T. (in press) The probability of conditionals: The psychological evidence. *Mind and Language.* [JSBTE]
Papineau, D. (1993) *Philosophical naturalism.* Blackwell. [rPC]
Paprafragou, A., Massey, C. & Gleitman, L. (2002) Shake, rattle, 'n roll: The representation of motion in language and cognition. *Cognition* 84:189–219. [rPC]
Parsons, T. (1990) *Events in the semantics of English.* MIT Press. [rPC]
Pawley, A. (1993) A language which defies description by ordinary means. *The role of theory in language description,* ed. W. A. Foley. Berlin: Mouton de Gruyter. [DIS]
Pelegrin, J. (1993) A framework for analyzing prehistoric stone tool manufacture and a tentative application of some early stone industries. In: *The use of tools by human and non-human primates,* ed. A. Berthelet & J. Chavaillon. Oxford University Press. [aPC]
Perner, J. (1991) *Understanding the representational mind.* MIT Press. [aPC]
(1998) The meta-intentional nature of executive functions and theory of mind. In: *Language and thought,* ed. P. Carruthers & J. Boucher. Cambridge University Press. [aPC]
Peterson, C. & Siegal, M. (1998) Representing inner worlds: Theory of mind in autistic, deaf and normal hearing children. *Psychological Science* 9:117–33. [aPC]
Pettersson, L. G., Blagovechtchenski, E., Perfiliev, S., Krasnochokova, E. & Lundberg, A. (2000) Recovery of food-taking in cats after lesions of the corticospinal (complete) and rubrospinal (complete and incomplete) tracts. *Neuroscience Research* 38(1):109–12. [LH-V]
Pietroski, P. (2003) The character of natural language semantics. In: *Epistemology of language,* ed. A. Barber. Oxford University Press. [rPC]
(2004) *Events and semantic architecture.* Oxford University Press. [rPC]
Pietroski, P. & Crain, S. (2001) Nature, nurture and universal grammar. *Linguistics and Philosophy* 24:139–86. [rPC]
Pinker, S. (1994) *The language instinct.* Penguin Press. [aPC, CM, PPS]
(1997) *How the mind works.* Penguin Press. [aPC, CP]
(2002) *The blank slate.* Viking Press. [RFB]
Place, U. T. (1956) Is consciousness a brain process? In: *The philosophy of mind: Classical problems/contemporary issues,* ed. B. Beakley & P. Ludlow. MIT Press. [PPS]
Pléh, Csaba. (2002) Kommunikationsmuster und kognitive Architekturen. In: *Allzeit zuhanden: Gemeinschaft und Erkentniss in Mobilzeitalter,* ed. K. Nyíri. Passagen Verlag. [CP]
Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso-Moyano, J., Myers, C. & Gluck, M. A. (2001) Interactive memory systems in the human brain. *Nature* 414(6863):546–50. [LH-V]
Popper, K. (1959) *The logic of scientific discovery.* Routledge. [rPC]
Povinelli, D. (2000) *Folk physics for apes.* Oxford University Press. [aPC]
Premack, D. (1986) *Gavagai! Or, the future history of the ape language controversy.* MIT Press. [aPC]
Pulvermuller, F. (2002) A brain perspective on language mechanisms; From discrete engrams to serial order. *Progress in Neurobiology* 574:1–27. [DB]
Pylyshyn, Z. (1973) What the mind's eye tells the mind's brain. *Psychological Bulletin* 80:1–24. [PPS]
(1978) Imagery and artificial intelligence. In: *Minnesota studies in the philosophy of science, vol. 9,* ed. C. W. Savage. University of Minnesota Press. [PPS]
(in press) Mental imagery: In search of a theory. *Behavioral and Brain Sciences.* [PPS, PC]
Quine, W. V. O. (1974) *The roots of reference.* Harvard University Press. [JLB]
Rao, R. P. N. & Ballard, D. H. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation* 9(4):721–63. [JJB]
Reber, A. S. (1993) *Implicit learning and tacit knowledge.* Oxford University Press. [JSBTE, PC]
Rey, G. (1997) *Contemporary philosophy of mind.* Blackwell. [rPC]
Rizzolatti, G., Fogassi, L. & Gallese, V. (2000). Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. In: *The new cognitive neurosciences, 2nd edition,* ed. M. S. Gazzaniga. MIT Press. [JJB, PC]
(2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience* 2:661–670 [GV]
Rubin, D. C. (1998) Beginnings of a theory of autobiographical remembering. In: *Autobiographical memory: Theoretical and applied perspectives,* ed. C. P. Thompson, D. J. Herrman, D. Bruce, J. D. Read, D. G. Payne & M. P. Toglia. Erlbaum. [RWS]
Rumelhart, D. (1977) Toward an interactive model of reading. In: *Attention and Performance VI,* ed. S. Dornic. Erlbaum. [RD]
Rumelhart, D. & McClelland, J. (1986) *Parallel distributed processing.* MIT Press. [rPC]
Ryle, G. (1968) A puzzling element in the notion of thinking. In: *Studies in the philosophy of thought and action,* ed. P. F. Strawson. Oxford University Press. [PPS]
Sacks, O. (1985) *The man who mistook his wife for a hat.* Picador. [aPC]
(1989) *Seeing voices.* Picador. [aPC]

Sanford, A. J. & Sturt, P. (2002) Depth of processing in language comprehension: Not noticing the evidence. *TRENDS in Cognitive Sciences* 6:382–86. [RD, PC]
Savage-Rumbaugh, S. & Lewin, R. (1994) *Kanzi, the ape at the brink of the human mind.* John Wiley. [aPC]
Schacter, D., Norman, K. A. & Koutstaal, W. (1998) The cognitive neuroscience of constructive memory. *Annual Review of Psychology* 49:289–318. [RWS]
Schaller, S. (1991) *A man without words.* Summit Books. [aPC]
Schrauf, R. W. (2000) Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture and Psychology* 6(4):387–417. [RWS]
(in press) A protocol analysis of retrieval in autobiographical memory. *International Journal of Bilingualism.* [RWS]
Schrauf, R. W. & Rubin, D. C. (1998) Bilingual autobiographical memory in older adult immigrants: A test of cognitive explanations of the reminiscence bump and the linguistic encoding of memories. *Journal of Memory and Language* 39(3):437–57. [SH, RWS]
(2000) Internal languages of retrieval: The bilingual encoding of memories for the personal past. *Memory and Cognition* 28(4):616–23. [SH, RWS]
(in press). On the bilingual's two sets of memories. In: *Connecting culture and memory: The development of an autobiographical self,* ed. R. Fivush & C. Haden. Erlbaum. [RWS]
Scribner, S. (1985) Vygotsky's use of history. In: *Culture, communication and cognition: Vygotskian perspectives,* ed. J. Wertsch. Cambridge University Press. [KN]
Segal, G. (1998) Representing representations. In: *Language and thought,* ed. P. Carruthers & J. Boucher. Cambridge University Press. [aPC, JGdV]
Shallice, T. (1988) *From neuropsychology to mental structure.* Cambridge University Press. [aPC]
Siegal, M. & Varley, R. (2002) Neural systems involved in "theory of mind." *Nature Reviews Neuroscience* 3:463–71. [GV]
Silverman, I., Choi, J., Mackewan, A., Fisher, M., Moro, J. & Olshansky, E. (2000) Evolved mechanisms underlying wayfinding: Further studies on the hunter-gatherer theory of spatial sex differences. *Evolution and Human Behavior* 21:201–13. [TW]
Slezak, P. (2002a) Thinking about thinking: Language, thought and introspection. *Language and Communication* 22:353–73. [PPS]
(2002b) The tripartite model of representation. *Philosophical Psychology* 15(3):239–70. [PPS]
(in press) Mental imagery: Déjà-vu all over again? (Commentary on Pylyshyn.) *Behavioral and Brain Sciences.* [PPS]
Slobin, D. I. (1996) From "thought and language" to "thinking for speaking." In: *Rethinking linguistic relativity,* ed. J. J. Gumperz & S. C. Levinson. Cambridge University Press. [DIS]
(2003) Language and thought online: Cognitive consequences of linguistic relativity. In: *Advances in the investigation of language and thought,* ed. D. Gentner & S. Goldin-Meadow. MIT Press. [DIS]
(in press) Language and thought on line: Cognitive consequences of linguistic relativity. In: *Advances in the investigation of language and thought,* ed. D. Gentner & S. Goldin-Meadow. MIT Press. [CM]
Smith, L. (2001) How domain-general processes may create domain-specific biases. In: *Language acquisition and conceptual development,* ed. M. Bowerman & S. Levinson. Cambridge University Press. [AC]
Smith, N. & Tsimpli, I-M. (1995) *The mind of a savant: Language-learning and modularity.* Blackwell. [aPC]
Smith, P. (1996) Language and the evolution of mind-reading. In: *Theories of theories of mind,* ed. P. Carruthers & P. Smith. Cambridge University Press. [rPC]
Sokolov, A. N. (1968) *Vnutrennaja recs i mislenyije.* Prosveschenyie. [CP]
(1972) *Inner speech and thought.* Plenum Press. [CP]
Sovrano, V. A., Bisazza, A. & Vallortigara, G. (2002) Modularity and spatial reorientation in a simple mind: Encoding of geometric and nongeometric properties of a spatial environment by fish. *Cognition* 85(2):B51–59. [LH-V, GV]
(2003) Modularity as a fish views it: Conjoining geometric and nongeometric information for spatial reorientation, *Journal of Experimental Psychology: Animal Behavior Processes,* in press. [GV]
Spelke, E. (1994) Initial knowledge: Six suggestions. *Cognition* 50:433–47. [aPC]
(in press) Developing knowledge of space: Core systems and new combinations. In: *Languages of the brain,* ed. S. Kosslyn & A. Galaburda. Harvard University Press. [aPC]
Spelke, E. & Tsivkin, S. (2001a) Initial knowledge and conceptual change: Space and number. In: *Language acquisition and conceptual development,* ed. M. Bowerman & S. Levinson. Cambridge University Press. [PR]
(2001b) Language and number: A bilingual training study. *Cognition* 57:45–88. [JJB, aPC, SH, PR]
Spelke, E., Vishton, P. & von Hofsten, C. (1995) Object perception, object-directed action, and physical knowledge in infancy. In: *The cognitive neurosciences,* ed. M. Gazzaniga. MIT Press. [aPC]
Sperber, D. (1994) The modularity of thought and the epidemiology of representations. In: *Mapping the mind,* ed. L. Hirschfeld & S. Gelman. Cambridge University Press. [SA]
(1996) *Explaining culture: A naturalistic approach.* Blackwell. [aPC, RS]
(2002) In defense of massive modularity. In: *Language, brain and cognitive development,* ed. I. Dupoux. MIT Press. [rPC]
Sperber, D., Premack, D. & Premack, A., eds. (1995) *Causal cognition.* Oxford University Press. [aPC]
Sperber, D. & Wilson, D. (1986) *Relevance: Communication and cognition.* Blackwell. (second edition, 1995.) [aPC]
(1995) *Relevance,* 2nd edition. Basil Blackwell. [JSBTE]
(2002) Pragmatics, modularity and mind-reading. *Mind and Language* 17:3–23. [rPC]
Spivey, M. J., Tyler, M. J., Eberhard, K. M. & Tanenhaus, M. K. (2001) Linguistically mediated visual search. *Psychological Science* 12:282–86. [RD]
Squire, L. R. (1992) Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review* 99:195–221. [RWS]
Stanovich, K. (1999) *Who is rational? Studies of individual differences in reasoning.* Erlbaum. [aPC, rPC, JSTBE]
Stanovich, K. E. & West, R. F. (2000) Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences* 23:645–726. [JSBTE]
Stein, E. (1996) *Without good reason.* Oxford University Press. [aPC]
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. & Sedivy, J. C. (1995) Integration of visual and linguistic information in spoken language comprehension. *Science* 268:1,632–34. [RD]
Tattersall, I. (2002) The case for saltational events in human evolution. In: *The speciation of modern Homo sapiens,* ed. T. J. Crow. Oxford University Press. [DB]
Thornton, R. (1996) Elicited production. In: *Methods for assessing children's syntax,* ed. D. McDaniel, C. McKee & H. Cairns. MIT Press. [rPC]
Tomasello, M. (1992) *First verbs: A case study of early grammatical development.* Cambridge University Press. [KN]
(1999) *The cultural origins of human cognition.* Harvard University Press. [DB, aPC]
Tomasello, M., ed. (1998) *The new psychology of language: Cognitive and functional approaches to language structure.* Erlbaum. [RD]
Tooby, J. & Cosmides, L. (1992) The psychological foundations of culture. In: *The adapted mind,* ed. J. Barkow, L. Cosmides & J. Tooby. Oxford University Press. [rPC]
Vallortigara, G., Zanforlin, M. & Pasti, G. (1990) Geometric modules in animal's spatial representations: A test with chicks. *Journal of Comparative Psychology* 104:248–54. [aPC, GV]
Varley, R. (1998) Aphasic language, aphasic thought. In: *Language and thought,* ed. P. Carruthers & J. Boucher. Cambridge University Press. [aPC, rPC, NC]
(2002) Science without grammar: Scientific reasoning in severe agrammatic aphasia. In: *The Cognitive basis of science,* ed. P. Carruthers, S. Stich & M. Siegal. Cambridge University Press. [aPC, NC, GV]
Varley, R. & Siegal, M. (2000) Evidence for cognition without grammar from causal reasoning and "theory of mind" in an agrammatic aphasic patient. *Current Biology* 10:723–26. [GV]
Varley, R., Siegal, M. & Want, S. (2001) Severe impairment in grammar does not preclude theory of mind. *Neurocase* 7:489–93. [GV]
Vygotsky, L. (1934/1986) *Thought and language.* Trans. Kozulin, MIT Press. [aPC, GO, KN, CP]
(1986) *Thought and language, revised edition.* MIT Press. [WF]
Walker, S. (1983) *Animal thought.* Routledge. [aPC]
Wason, P. & Evans, J. (1975) Dual processes in reasoning? *Cognition* 3:141–54. [aPC]
Watson, J. B. (1920) Is thinking merely the action of language mechanisms? *British Journal of Psychology* 11:87–104. [CP]
(1924) *Behaviorism.* Norton. [CP]
Welder, A. N. & Graham, S. A. (2001) The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development* 72:1653–73. [FX]
Wellman, H. (1990) *The child's theory of mind.* MIT Press. [aPC]
Wheeler, M. A., Stuss, D. T. & Tulving, E. (1997) Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin* 121(3):331–54. [RWS]
White, T. W. (2002) Unique and redundant connexin contributions to lens development. *Science* 295(5553):319–20. [LH-V]
Whorf, B. (1956) *Language, thought, and reality.* Wiley. [RFB, aPC, SH, CP]
Wierzbicka, A. (1985) The double life of a bilingual. In: *Polish People and Culture in Australia,* ed. R. Sussex & J. Zubrzucki. Australian National University Press. [SH]
Williams, D. M. & Hollan, J. D. (1981) The process of retrieval from very long-term memory. *Cognitive Science* 5:87–119. [RWS]

Williams, M. D. & Santos-Williams, S. (1980) Method for exploring retrieval processes using verbal protocols. In: *Attention and Performance VII,* ed. R. S. Nickerson. LEA. [RWS]

Wilson, T. (1985) Strangers to ourselves. In: *Attribution: Basic issues and applications,* ed. J. Harvey & G. Weary. Academic Press. [rPC]

(2002) *Strangers to ourselves.* Harvard University Press. [rPC]

Wilson, T., Hull, J. & Johnson, J. (1981) Awareness and self-perception: Verbal reports on internal states. *Journal of Personality and Social Psychology* 40:53–71. [rPC]

Wilson, T. & Stone, J. (1985) Limitations of self-knowledge. In: *Self, situations and social behavior,* ed. P. Shaver. Sage. [rPC]

Wittgenstein, L. (1921) *Tractatus logico-philosophicus.* Routledge. [aPC]

(1953) *Philosophical investigations.* Blackwell. [aPC]

Woolfe, T., Want, S. & Siegal, M. (2002) Signposts to development: Theory of mind in deaf children. *Child Development* 73:768–78. [rPC]

Wynn, K. (1990) Children's understanding of counting. *Cognition* 36:155–93. [aPC]

(1995) Origins of mathematical knowledge. *Mathematical Cognition* 1:35–60. [aPC]

Wynn, T. (2000) Symmetry and the evolution of the modular linguistic mind. In: *Evolution and the human mind,* ed. P. Carruthers & A. Chamberlain. Cambridge University Press. [aPC, TW]

(in press). Archaeology and cognitive evolution. *Behavioral and Brain Sciences.* [TW]

Xu, F. (2002) The role of language in acquiring object kind concepts in infancy. *Cognition* 85:223–50. [FX]

Zhang, S. & Schmitt, B. (1998) Language-dependent classification: The mental representation of classifiers in cognition, memory and ad evaluations. *Journal of Experimental Psychology: Applied* 4(4):375–85. [SH]