

Commentary on: P. Carruthers. The cognitive functions of language. *Behavioral and Brain Sciences*, in press.

Word Count: abstract = 60, text = 1382, references = 47, total = 1429

## A Metamodule for Conceptual Integration: Language or Theory of Mind ?

*Scott Atran*

*CNRS – Institut Jean Nicod  
1 Bis Avenue de Lowendal  
75007 Paris  
France  
Tel. 33-4 68 82 51 66 before 01-01-02*

*and*

*ISR – University of Michigan  
426 Thompson St.  
Ann Arbor, MI 48106-1248  
USA  
Tel. (734) 936-0458 after 01-01-02  
[satran@umich.edu](mailto:satran@umich.edu)*

Abstract. Those who assume domain-specificity, or conceptual modularity, must 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, and then integrate 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 begins with assuming 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 other) despite wide individual variation in exposure to diverse and fragmentary experiences (“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 (as I do), must contend with “Fodor’s Paradox,” or the problem of “combinatorial explosion.” For example, the fact that none of us find it difficult to hold and comprehend a sentence or thought in which, say, germs, gods and grease (or anything else, including the kitchen sink) coexist cannot owe to 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. For Sperber, this metamodule is part of folkpsychology, 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 1994:60). 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 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.”

Peter Carruthers offers a somewhat similar solution to “Fodor’s Paradox”; only, he attributes evolved metamodular functions to the language faculty rather than ToM. Carruthers’s reasoning and insights about the role of language in thought are enlightening (at least to me) and 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 also acknowledges that 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.

In fact, it is very difficult to see a way clear to one source or the other. This is because the crucial structural aspect of the metarepresentational system, embedding of mental states (in ToM), may have co-evolved with syntactic recursion (in language) (Atran in press). Short-term memory typically limits iterated embedding of mental states (e.g., “Peter thinks that [Dan knows that [...] to five or six levels (R. Dunbar); however, 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 (again, not predicted by associationist models, connectionist or other). Without this embedding process (whatever its source) whose surface product is “logical form” (in generative grammar’s technical sense), there seems no way to build cross-modular representations in sentences or thoughts.

In the end, preference for one source over the other may boil down to whose evolutionary stories one finds most congenial. But these stories (at least for the time being) are just speculative tales consistent with natural selection, lacking evidentiary standards that could rule out indefinitely many alternative and even contrary explanations. Although Carruthers aptly renders some of these stories as parenthetical remarks, his reasoning

suggests more to them than there really is. For example, when he says “this makes perfectly good-ecological-evolutionary sense” and leads to “just what one might predict,” there are, in fact, indefinitely many other sensible stories compatible with the data. At best, these accounts may retrodict findings. I find no case where they “predict” any surprising or significant discovery that was not available before the “evolutionary” account was proposed (this is a frequent problem with evolutionary psychology explanation, however interesting).

A couple of minor points are worth mentioning, although they do not affect Carruthers’s central claims. He may be too kind to those 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 doesn’t appear to have deep cognitive consequences (e.g., in reasoning, reference, pragmatic use or categorization). Moreover, the stimuli involved were routinely simple objects (cork, piece of paper). When more complex objects are involved, differences between Yukatek and English may disappear. For example, our team of anthropologists, psychologists and linguists who work with the Itza’ Maya (closely related ethnically and linguistically to Yukatek) found Itza’ more likely to group, say, a wood canoe (chem) with a metal motorboat (laanchaj) than with a tree (che’). Frequently cited differences in spatial reasoning (e.g., more cardinal than relative positioning because of lack of prepositions in the language) might 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).

Finally, Carruthers’s claim that the subtle social behavior of apes necessitates their being “capable of structured propositions of thought” seems unwarranted. There is no example I’m aware of indicating subject-predicate structures (i.e., propositions) in any creature save language-competent humans. Even chimps and bonobos consistently fail to apprehend such structures; novel “sentences” are maximally just two concatenated arguments with no subjects, such as “chase bite,” that humans shun (Atran & Lois 2001). If the claim were merely for communicating predicate-argument relations, without any argument being distinguished as the subject, there would be some independent support by analogy (although no direct empirical test or confirmation). For example, theories of a variety of forms of information representation (relational databases, formal logic, computer programming languages) and information processing (human vision, conceptual memory, real-time reasoning) hypothesize manipulation of predicate-argument relations (S. Pinker). A syntactic subject combines a logical function (a particular thematic role, typically agent) with the pragmatic function of topic in a topic-comment structure (allowing sentences to be pragmatically linked together in discourse); however, the only known case of an agent-focused thought (the logical-pragmatic subject) being structurally fit to a serial communication medium is human language.

Thanks to Noam Chomsky, Robin Dunbar, Steven Pinker for information cited.

Atran, S. (in press) Modest Adaptationism (commentary). *Behavioral and Brain Sciences*.

Atran, S. & Lois, X. (2001) Reply to Shanker. *Current Anthropology*, 42, 498-500.

Sperber, D. (1994) The modularity of thought and the epidemiology of representations. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind*. Cambridge University Press.