

## Cognitive Science and the Problem of Representation

Richard van Oort

*English, British Columbia*

**Abstract** In this article, I argue that the problem of representation stands at the center of the debate concerning the legitimacy of cognitivism as a research strategy for the humanities. Yet, curiously, very few commentators in this debate see representation to be a problem at all. Questions about the anthropological origin and function of representation tend to be regarded as at best supplemental, and at worst simply irrelevant, to the synchronic question of the causal mechanisms involved in the production of representation in the brain. I argue that this view is fundamentally mistaken and, furthermore, that we won't get clear about the central issues in the debate over cognitivism in literary studies until we get clear about the problem of representation. The problem is essentially one of how to define the human in terms of its most unique trait: the capacity for symbolic representation. After a review of how cognitivism misinterprets this question as a gradually evolved genetic adaptation of the nervous system, I turn to the theory of representation proposed by the neuroscientist and anthropological researcher Terrence Deacon in *The Symbolic Species: The Co-Evolution of Language and the Brain* (1997). Deacon stands out among cognitive scientists for his conclusion that language—symbolic reference—is an “evolutionary anomaly,” that is, inassimilable to the mechanism of gene replication. By understanding the exact nature of this evolutionary anomaly, we are in a much better position to assess the skepticism that is routinely directed toward those who use cognitive science to interpret literature. More precisely, I argue that the originary function of the sym-

I would like to thank my anonymous readers for their many helpful comments and criticisms and Meir Sternberg for his painstaking reports on earlier drafts and his meticulous editorial skills. I am also grateful for fellowship support from the Social Sciences and Humanities Research Council of Canada and the Critical Theory Emphasis at the University of California, Irvine.

*Poetics Today* 24:2 (Summer 2003). Copyright © 2003 by the Porter Institute for Poetics and Semiotics.

bolic sign is the deferral of lower-level indexical reference strategies. This originary anthropological function is most clearly evident in literature, art, and religion.

### 1. What Is the Problem of Representation?

Cognitive science has been making steady inroads into literary studies. Since the publication in 1980 of George Lakoff and Mark Johnson's influential *Metaphors We Live By*, literary critics have been encouraged by the idea of a cognitive poetics—of, that is, a systematic theory of the mind in which literature is not merely peripheral but central to the understanding of human psychology. Much of this cross-disciplinary work, however, has not produced the revolution in literary studies hoped for by its proponents.<sup>1</sup> On the contrary, cognitively informed interpretations of various literary works seem for the most part content to apply the newly acquired terminology of cognitive science to the fundamentally old task of providing original interpretations of literary works.<sup>2</sup> In this sense, cognitive poetics is new wine in old bottles.

Within literary studies, debates over the new cognitive paradigm, both pro and con, have largely avoided the bigger philosophical and anthropological issues. Indeed, one wonders if there is actually any difference in this debate other than the same old conflict about whose theoretical paradigm gets to be used when it comes to interpreting literature. Such debates are testimony to literature's continued (sacred) power over its interpreters.<sup>3</sup>

But I believe there are deeper issues at stake. In particular, I do not think

1. Fludernik et al. (1999: 395) predict that the “poetics . . . for the new millennium” will be a cognitive poetics. Richardson (1999) and Crane and Richardson (1999) review the situation and conclude that literary studies is lagging behind in its attention to the “cognitive revolution.” But undoubtedly the most outspoken commentator is Turner (1991). In his polemical introduction to *Reading Minds* (“Professing English in the Age of Cognitive Science”), he (ibid.: 3) claims that “contemporary critical theory . . . is ungrounded and fragmented” and derides it as a “self-sustaining” and “self-feeding” “mandarin activity,” “unrestrained by laws of entropy.” The point of this polemic seems entirely political, aimed at those discontented with “theory” in the humanities. This is unfortunate because it gives the impression that cognitive science is not likewise heavily theoretical. Turner paints a rather misleading picture in which the fresh, childlike wonder of cognitive science can reclaim the lost innocence of literary studies by removing it from the grip of a cynical and increasingly senile “critical theory” and returning it to the simpler, happier days of object perception and object manipulation.

2. This is the substance of the criticism of Fischer 1996; Gross 1997; and Jackson 2000.

3. In this respect, Gross (1997) is exemplary. Her criticism of Turner is motivated by the assumption that literature is somehow more powerful than cognitive science when it comes to explaining the human condition. But the real debate is not over whose interpretive strategy should be privileged in the exercise of interpreting literary texts, but what enables us to engage in this exercise in the first place. On this score, Turner at least has a hypothesis, whatever one may think of its scientific pretensions.

that the cognitive paradigm, at least as it is currently constituted, is ready to revolutionize the study of literature and culture. And this is ultimately because the cognitive model lacks a theory of representation adequate to the task of interpreting human culture.<sup>4</sup> Simply put, the cognitive paradigm has overlooked the anthropology implicit in its approach to the problem of the origin of representation. Or, to put the same point somewhat less charitably, it interprets this origin exclusively in terms of ontogenetic development, thereby failing to recognize that ontogeny can only offer a partial and limited contribution to the *specifically anthropological* problem of language origin and function.

What do I mean by representation? And why do I suggest that a specifically “anthropological” approach is needed for a theory of representation? In order to begin answering these two questions, it is necessary to ask what is assumed by anthropology. I take it for granted that anthropology depends upon biology, just as biology depends upon chemistry and physics. In making this assumption, I am not privileging one discipline over another. I am just saying that a lower level of explanation must be assumed by each higher-level discipline. For example, to explain the structure and function of organisms, the biologist assumes the existence of molecules out of which organisms are composed. The biologist’s job is not to explain the behavior of the elementary molecules (which is the job of the chemist), but rather to explain the behavior of the macromolecules—in particular RNA and DNA—that are capable of replication, which is to say, of life (see Dawkins 1989 [1976]). As Richard Dawkins and others (e.g., Daniel C. Dennett [1995]) have ably demonstrated, Darwin’s idea of natural selection is all that is necessary to explain the evolution of an immense variety of life-forms, given the reality of a mechanism for self-replication.

As biology stands to chemistry, so does anthropology to biology itself. If anthropology has the right to exist as a scientific discipline, it is because it can provide a systematic answer to questions that the biological level of explanation cannot. That such an answer can be provided is of course controversial. In anthropology (as in the social sciences and humanities more generally), the dispute between “culturalists” and scientists is really a dispute about whether anthropology is an interpretive discipline (Geertz 1973) or can be assimilated to the explanatory mechanisms of biology (Wilson 1978, 1999).

4. See Miers (1992), who is one of the few commentators to recognize that the major difficulty for what he calls the “new cognitivism” is its theory of representation. Jackson (2000: 330) remarks in his healthily skeptical review of cognitive approaches to literature that the dispute between poststructuralism and cognitive science properly begins with a theory of representation; however, he does not develop this notion.

So, how would a biologist explain the origin of representation? In biology, a trait shared by all members of a species can be explained from four logically independent perspectives, namely, those of mechanism, ontogeny, phylogeny, and function.<sup>5</sup> Mechanistic explanations address the causal mechanisms underlying the trait; for example, one can explain language in terms of the neurological, anatomical, and psychological structures necessary to produce speech. Ontogenetic explanations appeal to the genetic and environmental factors that guide the development of the trait in the individual from the embryo to adulthood, for example, in the stages of language acquisition by the child. Phylogenetic explanations trace the trait's evolutionary history by studying its appearance in (fossilized) ancestors. In the particular case of language, such studies are necessarily limited to speculations about the pattern of brain evolution in early hominids, such as *Homo erectus*, as well as to comparisons with other closely related living primates, such as chimpanzees and bonobos. Finally, functional explanation looks at the fitness consequences of the trait for survival and reproduction; in particular it offers hypotheses about the selection pressures necessary for the appearance and maintenance of the trait in question.

As Robin Dunbar (1998: 92) has argued, although these four perspectives are logically independent of one another, no explanation of a particular trait is complete without being considered from all four of them. We can now see that the particular emphasis of the cognitivist model of representation is on mechanistic and ontogenetic explanations. Neuroscience is making rapid advances in studying the structure of the brain and the brain's role in generating higher cognitive functions like language. Cognitive science takes its lead from neuroscience by studying the causal mechanisms involved in perception and conceptual thought. Cognitive science is also interested in ontogeny because, by studying the development of the individual, we get some idea of which cognitive structures are more basic and which require greater interaction with the outside (perceptual, social, and cultural) environment in order to develop. In the case of language, for example, it has long been clear that there is a "critical period" for language acquisition, so that both genetic and environmental factors play a crucial role. It is true of course that evolutionary psychology also promises to balance these well-developed areas of research by offering hypotheses concerning the evolutionary origin and function of higher cognitive functions like language (see, for example, Donald 1991; Pinker 1994): this takes us into the area of phylogenetic and functional explanation. But it is precisely in

5. See Dunbar 1998: 92–93 and Hauser 1996: 2 for applications of these four perspectives to language. See also Sebeok 1994: 39.

this area that the theoretical difficulty of an anthropological definition of language emerges most prominently.

The theoretical difficulty (as opposed to the obvious empirical difficulty involved in dealing with hypothetical rather than living specimens) is that one has to commit oneself to a definition of language in order to know what to look for. But this definition in turn depends upon an observation that is not merely empirical but indeed historical, which is to say, implied by the language user's participation in a linguistic community that ultimately extends back to the very first language users. To ignore this historical continuity in the hope of achieving greater objectivity in the definition is illusory, because it is precisely this historical continuity that defines language and differentiates it from the kind of genetically transmitted signal systems used by many other animals. But many are the scientists and their scientifically minded fellow travelers (see, for example, Englefield 1977) who have decided that language is fundamentally no different than other animal communication systems. It is this assumption that motivates researchers to claim, for example, that human language emerged from the kind of call system used by vervet monkeys (see Seyfarth et al. 1980).

On the other side of this debate stand the linguists, who for the most part remain committed to the anthropological specificity of language—to the point where, as in the case of Noam Chomsky, they deny that evolutionary theory can hope to explain the origin of language from animal signal systems. For the linguist, language is distinct from the kind of stimulus associations produced by conditioning or instinct. Alarm calls, such as those used by vervet monkeys, therefore do not qualify as evolutionary precursors because we now understand them to have a different structure and function. Whereas human language has a *symbolic* structure and function, animal signal systems have an *indexical* structure and function. More precisely, the referential function of an index is given by the interpreter's prior experience of the contiguous relationship between the indexical sign and its object. This relationship is most often perceptually learned (e.g., by stimulus/stimulus association or conditioning), but it can also be genetically assimilated, so that the perception of the referent will predispose the perceiver to utter the sign prior to any actual learning. Indeed, vervet monkey alarm calls appear to be genetically assimilated (i.e., "hardwired") examples of indexical signs (see Deacon 1997: 330–31). Symbols, on the other hand, do not work like this. Symbolic reference is not generated by empirical association between perceptual categories (whether learned or innate). Rather, a symbol is capable of generating a referential function even if we have never experienced the object that the symbol refers to. For example, I may never have tasted haggis before, but if it is described to me as "a large spheri-

cal sausage made of the liver, heart, and lungs of a sheep, all chopped and mixed with beef or mutton suet and oatmeal and seasoned with onion, cayenne pepper, and other spices" (*Encyclopaedia Britannica Online*), I get a pretty good idea of what haggis is without having the item directly before me where I can see, touch, and taste it. Indeed, when I come across an object that fits this description, I may readily identify it as haggis. Reference has then been produced not by directly associating the word *haggis* with a real haggis, but by producing a "clear and distinct" idea of the formal structure of haggis (i.e., haggis as a semantic category). In other words, reference to the object (i.e., the real haggis) is mediated by a network of other symbolic relationships (i.e., the description of haggis as a sausage made of lamb, oatmeal, onion, cayenne, and so on). Indeed, because symbolic reference seems to function only by combining symbols, we may decide that this combinatory feature is the distinguishing mark of language. Thus, only a system capable of syntax, including clause and phrase structure, qualifies as true language (see, for example, Bickerton 1990, 1995; Calvin and Bickerton 2000). In this case, representational systems that are symbolic but not syntactic, having a semantics but no well-defined grammar, are merely "protolanguages" or "pidgens" (Bickerton 1995: 28–29).

Like the linguists, I will be assuming that language can be most broadly defined as symbolic representation. For reasons that will become clearer later on, I also assume that symbolic representation implies syntax. I therefore see no need to distinguish so dramatically, as Derek Bickerton (1990, 1995) does, between protolanguage and language, between symbolic words and complex sentences. I would claim, with Terrence Deacon (1997), that once you have symbolic representation, you inevitably get syntax. Historically speaking, protolanguage inevitably leads to language, because the definition of *protolanguage* is that reference is generated symbolically, which means associating the sign with other signs in a minimal syntactical system *before* referring this system as a whole to a nonlinguistic object.

So much for the difference between mechanical, ontogenetic, phylogenetic, and functional explanations and for language as symbolic representation. But what makes an explanation specifically "anthropological"? It counts as such when its broadest level of application is limited to our own species. For example, if we assume that language is not a specifically anthropological phenomenon, but rather a general communication system with equivalents in the animal world, then we are making a biological claim about communication systems in general. Hence we will be inclined to interpret the function of language as we would any other biological trait, namely, in terms of its fitness consequences for the individual or, more precisely, for the individual's genes (Dawkins 1989 [1976]).

Ultimately, of course, the human world must answer to the biological world. No specifically anthropological category, such as language, is exempt from Darwinian processes of natural selection. But the peculiar difficulty in explaining the origin of a specifically anthropological institution, such as language, is that its *anthropological* function (as opposed to its more general biological function) cannot be reduced to the biological level of explanation without at once also evacuating it of its specifically anthropological function. Interpreted in terms of its biological function—its fitness consequences—language can only be seen as a more elaborate way of representing and manipulating the outside world. If I can say to you, “Be careful! There’s a bear behind that tree,” this saves you from having to discover the bear yourself by actually perceiving it with your own senses. Indeed, this view of representation is precisely what motivates ethologists to see language as simply a more advanced signal system that evolved from the kind of alarm calls used by many other animals.<sup>6</sup> I hope it is intuitively obvious how we can explain the origin and evolution of these types of call systems using only the mechanistic and functional explanations implied by natural selection. I will later elaborate more fully on the relationship between learned behavior and innate or genetically assimilated behavior, but for now it is only necessary to understand that the idea of language as a sophisticated signal system for drawing attention to objective features of the world can be quite easily assimilated to Darwinian principles.

What remain less tractable for this model of representation, however, are precisely those elements traditionally studied in the humanities, namely, literature, art, and religion. How are we to explain to our proverbial visitor from Mars the function of religious or aesthetic representations? It is in the interpretation of these forms of symbolic representation that we can discover no direct causal (genetic) relationship that would explain the fitness consequences of a particular cultural artifact.

Why is the interpretation of literature, art, and religion so intractable to functionalist and mechanistic biological explanation? Why do we always get the sense that something is missing, for example, when we are told that religious awe emerges in the individual’s natural experience of the might and power of the sun (Müller 1873: 366–69) or that “human culture is mainly a set of adaptations for courtship” (Miller 1999: 72)?

I think the reason for this skepticism is our sense that the causal relationship between the perception of worldly objects and their social-symbolic

6. For the original research involving vervet monkey alarm calls, see Seyfarth et al. 1980. For a synthetic review of the literature on animal communication, see Hauser 1996. For a critical assessment of the relation between vervet monkey calls and human language, see Burling 1993 and Deacon 1997.

representation in culture is not a simple one-to-one correlation. For example, as Emile Durkheim (1965 [1915]) realized in his criticism of Max Müller (1873), it is more likely that we experience awe toward the sun not because of its natural properties, the monotonous predictability of which will rather make the sun seem all but insignificant, but because of the sun's symbolic significance to a community of worshipers.<sup>7</sup> But this just begs the question of why we should not merely perceive the sun but also experience it as symbolically significant, as something worth worshipping collectively. It is this transition—from perception to symbolic representation—that needs to be explained. And it is precisely the functional explanation of this transition that cannot be reduced to purely biological terms. What we are seeking is rather an *anthropological* explanation, which applies only to those who can participate in the process of symbolic representation. Between the immediate perception of the symbolic object and its causal consequences for survival and reproduction lies a more complex representational relation, not to be transcended or reduced on pain of destroying the very anthropological function of the symbolic sign. It follows that an explanation of the symbolic relation must begin not with the reduction or elimination of this representational relation, but with a minimal account of its origin and function, an origin and function that is, furthermore, unique to our species.

To sum up the preliminary argument, I have defined the problem of representation and laid out some key terms. These key terms include: (1) the fourfold distinction among ontogeny, phylogeny, function, and mechanism in biological explanations; (2) the difference between biological and anthropological levels of explanation; and (3) the importance of a theory of representation for specifying higher-level anthropological categories. The assumption I am working with is that a definition of symbolic representation is also a minimal definition of the human and thus the starting point for an anthropology. As corollary of this assumption, such an anthropology, if it is to exist at all, must be distinct from—irreducible to—the explanatory categories of biology. Hence the explanation for symbolic representation must be in the first place a *functional* explanation, one moreover that recognizes the essential, specifically anthropological functions that symbolic representation performs. Such an anthropological hypothesis must be able to explain not merely instrumental uses of representation (e.g., how

7. As Durkheim (1965 [1915]: 103) shows, the “momentary impressions” that characterize perceptual experience “could not serve as a basis for these stable and permanent systems of ideas and practices which constitute religions.” Rather, “religion responds to quite another need than that of adapting ourselves to sensible objects” (ibid.: 102). The origin of religion therefore cannot be discovered in the “intrinsic properties” of objects: “The world of religious things is not one particular aspect of empirical nature; it is *superimposed upon it*” (ibid.: 261; Durkheim’s emphasis).



it enhances the individual's capacity to manipulate the environment), but also sacred and aesthetic uses of representation (e.g., how it constitutes our sense of moral reciprocity).

## 2. The Cognitive Model of Representation

Any reasonably complete account of representation must sooner or later address the problem of origins. Where do our representations come from, and how are they grounded in the real physical and biological world? In this section, I will look at two answers to this question: one given by Lakoff and Johnson (1980) and another, to be discussed at greater length, given by Mark Turner (1996). Though the two answers are not identical, they both adhere to the general assumption that understanding the causal relation between basic perception and higher cognitive functions, like language and culture (i.e., symbolic representation), is the best way to answer this question. As will quickly become clear, I think they take too much for granted in their assumption that there is a natural progression (evidenced in ontogenetic development) from basic perceptual categories to their *re*-presentation in language. Most damagingly, the assumption that there is such a progression leads them to reverse the causal order in their account of the origin of symbolic categories. Instead of understanding the construction of symbolic categories like language, ritual, and art as a consequence of a more general symbolizing function, they see symbolic categories as constructed by “emergent metaphors” (Lakoff and Johnson 1980: 58) or “projections” (Turner 1996: 17) from more basic perceptual and sensory experiences (e.g., tracking the path of a moving object, placing an object in a container, placing one object in front of another, etc.). But in so doing, they omit precisely what is peculiar about symbolic reference, namely, its “displacement” (Hockett and Ascher 1964; Burling 1993) or “decoupling” (Tooby and Cosmides 2001; Hernadi 2001) from the perceptual associations of direct experience. It is the originating process of displacement or decoupling that needs to be explained, not the subsequent recruitment of perceptual and sensory experiences for symbolic purposes.

Let me illustrate the problem of symbolic displacement or decoupling from an example given by Lakoff and Johnson (1980). In their discussion of metaphor, the authors inevitably get to the question of the “grounding problem,” that is, how metaphors refer not just to other metaphors, but to the real world as well.<sup>8</sup> To answer this question, they offer the following examples:

8. I take this example from an early work by Lakoff and Johnson, but the same problem recurs in, for example, Lakoff and Turner 1989: 113 and Lakoff and Johnson 1999: 45–59.

Harry is in the kitchen.  
 Harry is in the Elks.  
 Harry is in love. (Ibid.: 59)

According to Lakoff and Johnson, all three statements use the concept of a container. But in the first, the concept is not metaphoric but “emerges directly from spatial experience” (ibid.). Only the second and third statements are metaphors because they are dependent upon a mapping or projection from the prior *direct experience* of containment. Thus, the idea of containment is perceived in the first case, is mapped onto the idea of a social group in the second, and is mapped onto an emotion in the third.

But this explanation of how metaphors are grounded in the real world is really no explanation because it assumes precisely what is at issue, namely, the difference between a symbolic reference system and a reference system based on perceptual categorization. Thus, in those examples, a predicate refers to an increasingly abstract concept: first a room (the kitchen), then a social group (the Elks), and finally an emotion (love). Lakoff and Johnson appear to think that the fact that they have pointed out that the idea of a container gets steadily more abstract counts as an explanation of how metaphors are grounded. But in the first, crucial example, they fail to explain the difference between directly perceiving Harry in the kitchen and representing the fact that Harry is in the kitchen, for example, by saying “Harry is in the kitchen.” By assuming that the essential problem is fundamentally a problem of *individual* cognition rather than collective representation, they take for granted precisely what needs to be explained: namely, the social basis for higher cognitive processes like metaphor. Thus, they reduce the anthropological problem concerning the social basis of language to a cognitive problem concerning the mechanism and ontogeny of symbolic processes in the individual mind. But the problem of the social basis of symbolic representation—of language—is not simply supplemental to the cognitive structure of metaphor. On the contrary, it is the very origin of metaphor itself.

A second answer to the problem of the origin of representation is given by Turner (1996) in his study of “the literary mind.” Turner begins by making a series of broad and fundamental claims for his three central concepts: narrative, projection, and parable. He proposes, for example, that narrative is “the fundamental instrument of thought,” that it is a capacity “indispensable to human cognition generally,” and that it is our means of “looking into the future, of predicting, of planning, and of explaining” (ibid.: 4–5). The same fundamental status is claimed for parable, which arises “inevitably from the nature of our conceptual systems” (ibid.: 5). Turner is quite serious about the fundamental cognitive and biological status of narrative,

projection, and parable. At one point he even claims that the “motivations for parable are as strong as the motivations for color vision or sentence structure or the ability to hit a distant object with a stone” (ibid.).

These are strong claims indeed, but what exactly does Turner mean by narrative, projection, and parable? It turns out that these three concepts are ontogenetically and dialectically related. Narrative is the primitive or originary concept, because once you have the capacity to construct narratives, you can begin to project one narrative onto another. But by understanding one narrative in terms of another, you have not merely narrative, but parable. As a minimal example of parable, Turner (ibid.: 6) cites the well-known proverb, “When the cat’s away, the mice will play.” The parable works by projecting onto a specific situation (e.g., the behavior of a classroom of preschoolers without their teacher) an easily grasped story concerning our knowledge of the relationship between cats and mice. Cats control mice populations; when cats are not around, mice proliferate and cause havoc—they “play” just like preschoolers do without their teacher. We have projected one story, the story of the cats and mice, onto another story, the story of preschoolers without their teacher. Hence Turner’s (ibid.: 5–7) hierarchical triad: narrative → projection → parable.

But of course narrative, as Turner realizes, is not itself an atomic form. It depends on yet more basic processes of pattern recognition, including the ability to recognize objects and events. And the most basic patterns, Turner argues, are those perceptual and motor patterns by which we relate to the world around us. Thus, our cognitive processes appear to bottom out at the level of sensorimotor patterns. Turner calls these patterns “image schemas,”<sup>9</sup> such as perceiving milk flowing into a glass; drinking from a glass; recognizing the similarity between one door and another, one chair and another; sitting on a chair; opening a door; and so on. In other words, at the most basic level we form categories of objects and events. By recognizing the similarity between events, we create basic categories of events. Turner (ibid.: 17) argues that these sequences or “small spatial stories” are the basic elements of narrative. One could summarize Turner’s developmental schema of cognitive categories as follows: image schema → narrative → projection → parable.

It is important to remember that Turner regards this sequence as ontologically given, in the sense that, given our biological and evolutionary past, there is no way we can prevent our cognitive systems from producing narratives and projecting them to produce parables. This ontogenetic or developmental inevitability is no doubt a guiding assumption for research

9. The term is actually due to Mark Johnson (1987). See Turner 1996: 22–25 for a brief account of the main research.

in cognitive science. In a related discussion, for example, it enables Steven Pinker (1994) to talk of a “language instinct.” Turner’s idea of “the literary mind” is motivated by an equivalent faith in the cognitivist and innatist program. The difference is that, where Pinker argues for a biologically given language competence, Turner argues for a biologically given literary or aesthetic competence. Just as we are all language users, so too are we all poets, because we all possess an innately given literary or aesthetic capacity. In this sense, the separation of poetry from everyday language is a historical contingency rather than a cognitive necessity; the historical division has nothing to do with the cognitive reality that underpins our literary capabilities. When Turner claims that our motivation for parable is as strong as our motivation for color vision and sentence structure, he is putting forth a version of the strong innatist argument. In fact, he appears to go one better than Pinker, because he believes that language can be more parsimoniously explained as a product of narrative, projection, and parable (see Turner 1996: 140–68).

But Turner’s innatist argument for the literary mind suffers from the same problem as that for the linguistic mind: it takes the ontogenetic situation as paradigmatic of the evolutionary or phylogenetic situation. The argument for an innate language acquisition device (LAD) is that, because language emerges inevitably in the normal course of child development, despite the poverty of the available stimulus, it must therefore be an instinct or innate competence, hardwired from the beginning. In fact, I think that this argument, though mistaken, is actually *more* plausible than Turner’s hypothesis of an innate literary competence. Innatists like Steven Pinker (1994) and Derek Bickerton (1995) at least recognize that language is a species-specific trait, whose origin presents a unique challenge to evolutionary theory because there is simply no precedent for it on the biological record and hence no possibility of generalizing across species with similar traits. Turner, on the other hand, tends to downplay the evolutionary anomaly of the literary mind. He obviously believes that the literary mind is uniquely human (his second chapter is titled “Human Meaning”). However, in tracing the origin of human literary capacities to basic perceptual and motor functions, he opens himself to the charge that the literary mind is in fact not uniquely human but a universal feature of all creatures equipped with equivalent perceptual and motor capacities. The crucial “anthropological” question Turner avoids asking is why the presence of equivalent sensorimotor functions in many nonhuman species has not led to an equivalent literary or aesthetic capacity.

As I will show in my fourth section, this assumption gets Turner into

some hot water when it comes to confronting the evolutionary and anthropological implications of his argument. Now I would like to comment on an internal difficulty that emerges in the course of his own discussion and reflects the larger anthropological issues that I wish to emphasize.

Turner, like Lakoff and Johnson, implies that what he calls “bodily” or “spatial” stories are likely primary and that more abstract stories about thought and social reality are projections from these basic sensorimotor patterns. Yet, duly cautious and empirical, he suggests that not all the evidence is in yet, so we cannot be sure if it is absolutely true. For example, the claim that “abstract thought and reasoning are always grounded . . . in spatial and bodily stories,” although “not clearly false,” is nonetheless “too extreme for the available evidence” (Turner 1996: 51). The most we can assert, Turner (*ibid.*) concludes, is that such a claim is “plausible.”

Why does Turner shy away from stating outright that abstract thought is based on the projection of direct sensorimotor experience? What kind of evidence other than that he depends on for his entire analysis—the evidence of our everyday linguistic metaphors—could persuade him that his claim is not merely plausible but true? Why the hesitation?

I think the reason for this hesitation is very significant, because it demonstrates the realization that metaphor or projection is not a one-way street from simpler levels of direct, unmediated bodily experience to more complex levels of indirect, symbolically mediated experience, and this fact throws into question Turner’s basic assumption that narrative projection is a continuous development from lower neurological processes to higher, more abstract ones. That is, Turner notices that, although most of our expressions do seem to project from spatial or bodily experience to more abstract concepts, projection can also work the other way around. So, for example, if my car fails to start, and I say, “It appears my car has different ideas about whether I make the reception,” I am projecting onto the event of the car’s not starting a story of disagreement between partners in a dialogue. Rather than projecting “upward,” so to speak, from object to mind (e.g., “My promise to make the reception is now *broken*”), I am projecting “downward” from mind to object. I am imposing on a physical object (the car) the state of being in which it has a mind capable of ideas and, in particular, of disagreeing with other language users.

It is important to see exactly what is at issue here. The difficulty lies not in the fact that metaphors work both ways—from body to mind and from mind to body, so to speak—but in the conclusion that, because so many of our metaphors are body- and space-oriented, therefore metaphor *must itself be describable in terms of the same physical biological process* by which we pro-

duce associations between perceptions and motor responses. The conclusion simply does not follow, because what is really at issue is not the presence of visual and manual biases in metaphoric processing, but the origin of metaphor as such. By explaining this origin in terms of more basic sensorimotor functions, Turner interprets the presence of an inherent cognitive bias as sufficient evidence for an account of the origin of a referential process. But it is the anthropological structure and function of this referential process, not the inherent bias toward spatial and bodily reasoning, which needs to be explained.

What is so unique about the referential process of metaphor? Turner, Lakoff, and Johnson locate the uniqueness in the fact of “embodied experience,” that is, of having a human body capable of interacting with the outside world. I agree that this is indeed a necessary condition of metaphor, but it is a necessary condition in the same sense that eating, sleeping, and reproducing are necessary for metaphor. In other words, it does not really address the specific nature of the problem of metaphoric processing. What is in principle one of the strengths of the cognitivist paradigm, namely, its commitment to general scientific categories of explanation, becomes a weakness when the ambition for a general explanation is stretched too far and, in the process, erases precisely what is distinctively human about metaphoric thought.

What, then, is so unique about metaphor, narrative, and other fundamental anthropological categories? I believe that the problem of their origin is really a problem of referential processes in general and particularly of how words or symbolic signs relate to the world. It is therefore a problem inseparable from the problem of language, which is to say, of symbolic reference generally. To tackle it, we therefore need to rethink the cognitive processes involved in acquiring and maintaining symbolic reference. Buoyed by advances in neuroscience and computer modeling, cognitive science is eager to interpret symbolic referential processes in terms of more basic neurological processes. But I think we need to be more cautious before rushing to draw parallels between symbolic functions and general sensorimotor functions. Until we have appreciated the specific problem of symbolic reference, we are ill positioned to grasp the causal relationships in the evolution of language and the brain.

### 3. Symbolic Reference

In this section, I will review the reasons why symbolic reference cannot be so easily assimilated to basic perceptual and motor functions. This argument will refer in some detail to Terrence Deacon’s *The Symbolic Species* (1997), a

wide-ranging discussion of the origin of symbolic representation by a neuroscientist and evolutionary anthropologist.<sup>10</sup>

First, however, a word about terminology. By *symbolic reference* or *symbolic representation* or *symbolic functions*, I mean the general semiotic process of interpreting symbolic relationships. In this sense, language is of course symbolic, but so are other specifically human institutions, such as rituals, games, literature, music, and mathematics. In fact, I think that symbolic reference is the general category by which to understand other specifically human phenomena (e.g., language, art, and ritual). Insofar as metaphor (Lakoff and Johnson 1980) and narrative (Turner 1996) are also fundamental human categories (and I believe they are), then they too must be understood as constituted by the process of symbolic reference. The difficulty is that symbolic reference is not an easy category to define. Indeed, according to Deacon (1997: 34), it is an “evolutionary anomaly” that cannot be readily assimilated to presymbolic neural and cognitive precursors.

For Deacon, given this “anomaly,” the main problem for a general theory of human cognition is how you can get from an indexical reference system to a symbolic reference system.<sup>11</sup> This transition cannot be explained exclusively by neurobiology but requires, in his view, a more global socioecological perspective. Deacon’s theory is quite breathtaking in its scope and application. By summarizing the basic theoretical argument, I hope to give some idea of the significance of his anthropological research for students of literature and aesthetics.

The key lies in his model of cognition as fundamentally a process of interpreting increasingly complex layers of reference. In an illuminating application of C. S. Peirce’s semiotic categories of icon, index, and symbol, Deacon describes each mode of reference as requiring a different level of cognitive

10. In general, Deacon’s book has been very well received by the scientific community. See, for example, reviews in Calvin 1997, Holloway 1998, Hurford 1998, Hudson 1998, Ruse 1998, Turner 1998. From the humanities, on the other hand, there has been little interest in Deacon thus far; for a notable exception, see Gans 1999. Given my criticism of Turner’s *The Literary Mind*, the reader may be surprised to see his name in the above list. But Turner’s brief review, though very favorable, seems not to notice the radicality of Deacon’s argument that human neurobiology has been shaped by something as abstract as a word.

11. Deacon is certainly not the first to put the problem in this way, but I think his account represents the best recent synthesis of evolutionary neuroscience and semiotic theory. Similar arguments for the anthropological specificity of the difference between the sign-as-index and the sign-as-symbol can be found in Cassirer 1944: 24–25; White 1949: 22–39, 1987: 259–72; and Bickerton 1995. Also noteworthy is Holloway’s (1969: 395) insightful analysis of hominid toolmaking as evidence for the origin of the specific difference of human cognition, in that it requires the symbolic process of the “imposition of arbitrary form on the environment.” See also Burling 1993 for an excellent discussion of the difference between the “digital” linguistic system and the “analogue” primate call system. For the contrary argument that no essential difference between human language and animal communication exists, see Englefield 1977.

response. Cognition of the reference type is hierarchical in that each level implies the lower levels of reference. Thus, for example, indexical reference is not possible without first being able to recognize objects iconically.

Whereas Peirce (1956 [1897, 1903]) originally defined his three categories according to a general semiotic in which icons, indices, and symbols were seen as different aspects of an *already existing* sign system (see Sebeok 1994: 68), Deacon in effect anthropologizes Peirce's triadic schema. He argues that iconic and indexical representational processes are supported by all nervous systems, irrespective of their size and complexity, but symbolic representation has evolved in only one species: our own. Deacon's use of Peirce's semiotic categories of icon, index, and symbol may strike some readers as unorthodox.<sup>12</sup> But I think his reinterpretation builds very constructively on some of Peirce's basic insights. In summarizing Deacon's basic theory, therefore, I will try to show not its faithfulness to Peirce but how it furthers the debate concerning the origin of specifically human cognitive capacities.

How does Deacon explain iconic and indexical reference? At the most elementary level, iconic reference is equivalent to stimulus generalization. It is the default mode of representation, when we recognize the similarity between one stimulus and another. Iconic reference is not given by the inherent similarity between two objects; it is produced by an act of cognition that takes place *before* we even understand that what we have are two distinct objects that resemble one another. "The interpretive step that establishes an iconic relationship," Deacon (1997: 74) argues, "is essentially prior to this [distinction], and it is something negative, something that we don't do. It is, so to speak, the act of *not* making a distinction." Deacon's point here, which admittedly seems extremely counterintuitive and certainly not in line with Peirce's original schema, is that iconic reference is the most general reference type available to any creature equipped with a nervous system. In order to make predictions about the world, you need to generalize from particular cases. But in order to generalize, you need to have an interpretive system that produces not a distinction between discrete icons — this would assume the possibility of indexical reference — but resemblance itself. Only

12. For examples of the confusion that Deacon's evolutionary application of Peirce's semiotic categories has caused among linguists, see Hurford 1998 and Hudson 2001. I think the confusion points to a genuine difference between a metaphysics that takes symbolic reference for granted, as linguistics in general does, and an evolutionary perspective that attempts to ground symbolic reference in more basic referential processes. For example, what Deacon calls "iconic reference" may seem unrecognizable to followers of Peirce, because he grounds this referential process in basic categorization. The beauty of Deacon's solution is that it puts the evolutionary problem of human origin in terms of a semiotic processing problem.



after the act of *not* making a distinction—of noticing *only* likeness or sameness—can one decide, upon further reflection, that what one is observing is in fact an iconic relationship between two *distinct* objects. From the perspective of indexical reference, iconic reference is thus a failure to observe difference. Like the failure to notice the polar bear blending perfectly into its snow-covered environment, all one sees is white on a background of more white.

Deacon's analysis of iconic reference may seem somewhat idiosyncratic.<sup>13</sup> Isn't an icon defined by its resemblance to its object? For example, a portrait of George Bush is iconic because we recognize in the portrait features that we recognize in the real George Bush. But that is precisely the point! What inclines us to see the portrait under a primarily iconic function is the fact that so many features that we see in the real George Bush, we also see

13. Deacon has been taken to task for his interpretation of the icon. For example, Hudson (2001: 130) suggests that Deacon is guilty of reducing the iconic sign to basic categorization, which is clearly not part of any semiotic system. But I think Hudson has missed the fundamental point of Deacon's analysis, which is to explain the cognitive basis of the icon. Whether we are talking about a picture (the classic example of an icon) or the formation of basic categories, the role of iconic recognition is fundamentally the same. The difference is that in interpreting a picture we are involved in a complex hierarchy of iconic, indexical, and symbolic relationships. Within this hierarchy of relationships, we emphasize the iconic aspect by seeing the resemblance between the subject matter of the picture and the real or fictional object it refers to. We could just as easily stress the indexical relationship between the picture and its immediate surroundings (e.g., the fact that the picture is framed and hanging on a wall indicates that the picture is an art object of some value) or the symbolic relationship (e.g., the picture as a metaphor of social malaise in the twentieth century). I think, furthermore, that Deacon's analysis is in the spirit of Peirce's (1956: 104) original definition of the icon as a sign "whose Representative Quality is a Firstness," where the categories of "Firstness," "Secondness," and "Thirdness" imply the ascending hierarchy of semiotic relationships from icon to symbol. (Note, however, that I do not mean to imply that there is a simple correspondence between Peirce's categories of firstness, secondness, and thirdness and icons, indices, and symbols.) A sign "by Firstness," Peirce (*ibid.*: 104–5) argues, "can only have a similar object." Similarity is fundamentally a resemblance between the qualities of the sign and the qualities of the object, such that they "excite analogous sensations in the mind" (*ibid.*: 114). In the elementary case of "Firstness" (to use Peirce's terminology), similarity or iconicity is defined by a lack of negation or contrast, what Peirce calls "Secondness" (*ibid.*: 105). But, of course, in order to notice the similarity between distinct sets of objects, one must have already negated the first object by reflecting on its resemblance to another (slightly different, but nonetheless similar) object. Stimulus generalization is the application of one stimulus to another in order to do inferential work. But as Deacon rightly argues, we should not forget that such an inferential application also depends upon something we do not do, which is the act of not making *any* distinction. As Deacon (1997: 74) puts it: "Usually, people explain icons in terms of some respect or other in which two things are alike. But the resemblance doesn't produce the iconicity. Only *after* we recognize an iconic relationship can we say exactly what we saw in common, and sometimes not even then. The interpretive step that establishes an iconic relationship is essentially prior to this, and it is something negative, something that we don't do. It is, so to speak, the act of *not* making a distinction."

in the portrait. Naturally, we also know that the portrait is not identical with George Bush, but our ability to notice the *difference* between the portrait and the real George Bush is not a function of iconic interpretation itself. It is instead a function of our capacity to bring into correlation *higher levels* of iconic recognition, which we can only bring into focus if capable of distinguishing between sets of icons. In admiring the subject matter of the portrait, for example, we do not concentrate on the wooden frame, or the dust, or the reflection of the light: all these are examples of iconic recognition that we bracket in our experience of the portrait. We recognize the similarity between this frame and other frames, this wood and other wood, this dust and other dust, this reflection and other reflections, but in appreciating the portrait, we forget about these other features and concentrate on the likeness between the picture and George Bush.

For how iconic reference functions at this most basic level, compare Deacon's (1997: 75) own example of a camouflaged moth, whose wing coloration matches the color of the bark of a tree. A bird scanning the bark for prey will *not* see the moth if it notices *only* the resemblance between the moth and the tree. If the moth moves, or if some flaw in the moth's coloration alerts the bird to the *difference* between the bark and the moth, then the iconic interpretation is broken and the bird is forced to reassess its perception of the moth as just more bark. If the bird is alert, it may see this difference as an *index* of a moth.

This example matches our iconic interpretation of the portrait. Naturally, the latter is more complex because we are dealing with many more variables. To notice the moth, the bird needs to recognize the difference between just two categories: bark and moth. If its iconic interpretation applies to these two categories (bark and moth), then it is fine shape. If it only notices one category (bark), then it is the moth that is in fine shape. The same applies to the portrait. If our iconic interpretation is so general, or so limited, that we only notice the categories of wood, paint, and canvas, then we will see no George Bush in the portrait. If, however, we are attuned to facial recognition, as indeed humans are, then the face will appear immediately. But the fact that we also notice the wood and the canvas means, precisely, that we also know that this is not really George Bush. Our iconic recognition processes stop short of stimulating us to reach out and shake George Bush's hand. (The thrill of seeing lifelike wax figures of celebrities depends largely on the shock value of feeling as if one really were in the presence of someone famous like George Bush.)

Indexical reference builds on iconic reference. In order to pick out the moth against the background of the bark, or to notice the polar bear against its snowy background, one must notice a clue or *index* that points us toward

the presence of something other than just more bark or just more snow. If one were to locate the difference between the polar bear and its white background in, for example, its black nose, then one would have an index of this difference. Instead of just seeing more snow (e.g., white, white, white, white . . .), we notice a difference (e.g., white, white, white, not-white . . .). Perhaps after some practice, one might become adept at noticing polar bears, in which case the terrain becomes not just an endless white expanse, but also a varied environment with a lot of different clues indicating the presence or absence of polar bears (e.g., smells, tracks, movement, startled seabirds, and so on). Indexical reference entails the ability to interpret how one stimulus can *point to* or *indicate* another based on a pattern of predictable co-occurrence between distinct stimuli, where each stimulus is in turn interpreted iconically. Deacon (1997: 78) points out that iconic and indexical references are similar to the more traditional categories of perceptual generalization and learned association. What he wants to preserve in his notion of kinds of reference, however, is the fact that these are essentially representational strategies that increase the predictive or inferential powers of cognitively aware organisms. Icons and indices allow external physical relationships to be represented internally, which in turn allows a creature to adapt more readily to a changing environment. Iconic and indexical processes are thus indispensable features for cognitively based adaptation and evolution.

But if icons and indices are instances of an economy between internal representation and external reality, then it becomes possible for such internal representational processes to become the basis for external communication processes. And indeed, indexical communication systems abound in nature. The dog's snarl, the peacock's tail, the rabbit's thump—all are examples of communication, where a particular stimulus is associated with a particular behavioral response: aggression in the case of the dog, mating opportunity in the case of the peacock, danger in the case of the rabbit. In fact, some indexical communication systems seem to demonstrate the same level of abstraction that linguists have traditionally reserved for language. For example, vervet monkey calls are indexical signs that have evolved to refer to specific categories of objects, for example, categories of predators, such as leopards, snakes, and eagles. Superficially, vervet alarm calls appear to be equivalent to language, because an arbitrary sign (the call) refers to an object (the particular predator). But, Deacon argues, arbitrariness is not a sufficient condition for symbolic reference. Such a viewpoint fails to grasp the crucial difference between indexical and symbolic reference.

This difference is a consequence of the multilevel hierarchy involved in the move from indexical (and iconic) to symbolic relationships. Just as

indexical reference requires the ability to organize lower-level iconic relationships, so symbolic reference requires the ability to organize lower-level indexical and, ultimately, lower-level iconic relationships. But the specific difficulty in moving from index to symbol is that “the added relationship” required by symbols is “not a mere [empirical] correlation” between the sign and its object (Deacon 1997: 78). It bears, rather, on the preexisting formal relationship between signs. A word does not refer directly to an object or referent. Instead, reference is mediated by the word’s sense or meaning, which in turn is given by associating the word, not with the object, but with other words and other meanings. When we look a word up in the dictionary, we are putting this theory into practice. We understand a new word not by actually gaining physical access to a real-world object or referent, but by associating it with other words, that is, with other meanings. As Deacon (1997: 70) notes, the “correspondence between words and objects is a secondary relationship, subordinate to a web of associative relationships of a quite different sort, which even allows us reference to impossible things.”

There remains the difficulty, however, of developing a plausible hypothesis for the origin of symbolic reference given the preexisting indexical system. From an originary anthropological perspective, Deacon’s interest in Peirce’s semiotic categories of icon, index, and symbol focuses on the difference between the last two. In particular, the radically different cognitive strategy required by symbolic reference poses unique challenges to all non-human species, including our presymbolic ancestors, whose communication systems are based on indexical reference. The ubiquitous presence of the latter in the animal world suggests that symbolic representation originated as a response to a radically *different* set of selection pressures than those affecting the biological evolution of indexical systems. Only by understanding the different cognitive demands of symbolic reference, Deacon (1997: 34) argues, can we appreciate the “evolutionary anomaly” of language and why other species have failed to acquire it.<sup>14</sup>

14. Deacon notes that the tendency to interpret animal communication systems as simply less-advanced forms of language, though superficially plausible, is actually perverse in the extreme because it implicitly treats all forms of communication as exceptions to what is in fact itself an exception, namely, language. Researchers who are eager to see in monkey alarm calls the rudimentary nouns and phrases of full-blown language should take note of what Deacon (1997: 52) says on this score: “One reason we have such a difficulty [i.e., in distinguishing between signals learned by rote and true symbolic understanding] is that we don’t know how to talk about communication apart from language. We look for the analogues of words and phrases in animal calls, we inquire about whether gestures have meanings, we consider the presence of combination and sequencing of calls and gestures as indicating primitive syntax. On the surface this might seem to be just an extension of the comparative method: looking

Reconsider this difference. In order to be interpreted indexically, a sign must be associated either spatially or temporally with its referent. Thus, for a stimulus, such as a red light, to function as a sign and refer to an object or event, such as the dispensing of food from a chute, the stimulus must be habitually associated with the object or event. If the sign/object association is not maintained—if, for example, the red light were to go on over repeated trials without food being dispensed—then it would rapidly be forgotten. Notice that this applies also to “innate” signs, such as vervet monkey calls. If all vervet predators were removed from the monkeys’ habitat, then there would be no selection pressure to maintain the indexical link between sign and object; the signs would either disappear altogether or—perhaps more likely—would be adapted to another purpose for which there *was* selection pressure.

This is not the case with symbolic reference. We remember the meaning of a word independently of the appearance of its worldly referents, which may never be present (as in the case of God). This is because symbolic reference is generated by the relations between words rather than by the association between an individual word and its object. Only by understanding the virtual system *first* can we then agree on its corresponding reference. As Deacon (1997: 83) puts it, language implies a fundamental “duality of reference” because it refers only by delaying or negating reference to the world in favor of reference to other words. A word (as all Saussureans know) relates primarily to other words. Reference to a nonlinguistic world is still possible, of course (symbols still depend upon lower-level indexical and iconic processes), but it is now pursued on a different basis than the kind of one-to-one associations that characterize the indexical link between the sign and its object. What distinguishes symbolic from indexical reference is the mediation of the indexical object or referent by the relationships between signs themselves.

It is at this point that a theory of the *origin* of symbolic reference runs into a paradox. If words imply a dual reference, both to other words and to nonlinguistic objects, how can a symbol system ever get off the ground in the first place? Failing other words or symbols to refer to, the only type of reference available to the first word or symbol would still be indexical. How can an indexical sign ever become a symbolic sign?

Deacon’s solution to this problem appears complex because he ap-

---

for the evolutionary antecedents of these language features. But there is a serious problem with using language as the model for analyzing other species’ communication in hindsight. It leads us to treat every other form of communication as exceptions to a rule based on the one most exceptional and divergent case. No analytic method could be more perverse.”

proaches it from many angles, including a central discussion of the evolution and expansion of the prefrontal cortex in hominid evolution. But the basic theory is, I think, quite straightforward. Symbolic reference requires a shift in overall cognitive perspective. This shift occurs when a group of independently learned indexical signs are understood as bearing a relationship not merely to their co-occurring associated objects, but to one another. Perceiving the priority of the sign/sign relationships over individual sign/object relationships is the key to symbolic reference. But how is the shift in perspective produced?

To illustrate this problem, Deacon (1997: 84–92) cites the language-training experiments conducted by the primatologist Sue Savage-Rumbaugh (1986). The experiments were designed to teach two chimpanzees, Austin and Sherman, how to use a true symbolic, as well as an indexical, sign system. To do this, Deacon (*ibid.*: 84) argues, the chimps would have to refer arbitrary symbols or “lexigrams” (as the experimenters called them) not merely to the associated objects, but to other symbols or lexigrams: the salient test for whether or not the chimps had acquired a symbolic system is whether they could learn “to use lexigrams in combinations (e.g., syntactic relationships).”

In order to test whether the chimps were capable of learning this, the experimenters began with an extremely simple symbol system, involving just two “verbs” (lexigrams glossed as *give* and *pour*) and four “nouns” (lexigrams for different food or drink objects, i.e., *bananas*, *beancake*, *juice*, and *milk*). Thus, the chimps had to learn not merely the one-to-one indexical relationship between each food and drink item and its corresponding lexigram, but also the correct combination of lexigrams. For example, in order to receive juice, the chimp would have to produce the combination “pour juice”; in order to receive banana, “give banana.” Sequences such as “juice banana” or “juice pour” or “pour banana” or “give banana beancake” would not be rewarded.

From the naïve chimpanzee perspective, it is not at all obvious which sequences are going to be successful: “Even with this ultra-simple symbol system of six lexigrams and a two-lexigram combinatorial grammar, the chimpanzee is faced with the possibility of sorting among 720 possible ordered sequences ( $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ ) or 64 ordered pairs” (Deacon 1997: 85).<sup>15</sup> Given no previous experience or knowledge of even so simple a grammar, how can the chimpanzee learn which relationships are allowed and which excluded? To solve this problem, the experimenters engaged in an elaborate

15. According to my calculations, there are only thirty-six possible ordered pairs. But Deacon’s point still stands.

training program, in which the chimps were taught to distinguish between correct and incorrect sequences. In some cases, ingenious devices, such as beeping food dispensers (Savage-Rumbaugh 1986: 93–94), were introduced to get the chimps to notice when a correct sequence did not produce the desired result (e.g., the chimp requested a banana correctly, but the banana dispenser was empty—the point being to draw attention to the empty dispenser so that the chimp could modify the request, e.g., by requesting a different food item). Eventually, the chimpanzees were able to produce the correct sequences with the correct reference every time. Had they learned an elementary syntax? Had they learned that reference was a function not of indexical association, but of the relationships between the lexigrams themselves? To test for this, new food and drink objects, together with the corresponding lexigrams, were introduced into the existing system. If the chimps responded by correctly using the new lexigrams, then it could reasonably be assumed that they were using these new lexigrams as components of a symbolic system. Indeed, Austin and Sherman did use the new lexigrams appropriately, with very few errors or none. What accounts for this remarkable rise in success? What enabled the chimpanzees to go from producing correct combinations only after hundreds or thousands of trials to producing correct combinations with new lexigrams either immediately or with very few errors?

The answer, Deacon argues, comes from the different learning and memory strategy enabled by the possibility of combining lexigrams in an elementary syntax. Faced with the task of remembering which lexigram combinations are correlated with which features of the objective world, the chimps were presented with a choice: remember the individual correlations between each lexigram and the world (i.e., by essentially treating the lexigrams as additional stimuli to keep track of in an environment filled with competing stimuli) or, alternatively, use the lexigrams themselves as a more economical means to encode their understanding of how not merely to keep track of, but to *predict* regularities in the world. In other words, instead of being distracted by all kinds of irrelevant features of the world (e.g., the positioning of the lexigrams on the keyboard, the previous successful lexigram combination, preferences for a particular kind of food, etc.), the chimps learned that all these variables could be “off-loaded” (so to speak) and that a systematic set of relationships between the lexigrams could be substituted in their place. These relationships constitute a set of logical rules of inclusion and exclusion, which allow only two-word combinations of a “verb” (requesting a liquid or solid) and a “noun” (naming a particular food object). When new food items are introduced, their lexigrams can be easily integrated into this system. As Deacon (1997: 86) puts it, in shifting to a symbolic

mode of interpretation, the chimps “had discovered that the relationship that a lexigram has to an object *is a function of* the relationship it has to other lexigrams, not just a function of the correlated appearance of both lexigram and object.”

The experiment challenges the notion that language originates by way of a (highly unlikely) genetic mutation. Unlike modern *Homo sapiens*, Austin and Sherman are *not* genetically predisposed to interpret perceptual or indexical relationships in terms of a more economical system involving minimal symbolic signs. In learning their elementary symbol system, they nonetheless acquired the ability to generalize not simply on the basis of perceptual experience or conditioned association, but also on the basis of logical or categorical relationships between symbols or lexigrams. To be sure, the relationship between one lexigram and another is a form of association, in the sense that the chimps were trained to expect a “noun” to be paired with a “verb.” But the point is that the associations between lexigrams are *a more abstract and more economical way* of encoding the indexical associations or correlations between real objects and events. The key to symbolic reference lies in the ability to project the relationships existing between easily producible symbols *onto* features of the perceptual world. As Deacon (1997: 86–88) puts it:

In the minimalistic symbol system first learned by Sherman and Austin, reference to objects is a collective function of relative position within this token-token reference system. No individual lexigram determines its own reference. Reference emerges from the hierarchic relationship *between* these two levels of indexicality, and by virtue of recognizing an abstract correspondence between the system of relationships between objects and the system of relationships between lexigrams. In a sense, it is the recognition of an iconic relationship between two systems of indices. Although indexical reference of tokens to objects is maintained in the transition to symbolic reference, it is no longer determined by or dependent on any physical correlation between token and object.

The experiment is illuminating because it highlights the specific learning and memory difficulties presented by conceptualizing symbolic relationships. The ability to see the higher-order combinatorial regularity behind a logically related system of symbols is far from a natural or obvious step for the average chimpanzee. The major initial obstacle for the chimps was simply the bias produced by the indexical learning strategy by which the chimps learned the symbols in the first place. Given enough trials (and these often ran into the thousands), the chimps could learn the individual indexical symbol/object relationships. But the symbol/symbol relationships proved more difficult, because understanding the logic behind this higher-order system of relationships required abandoning their previous



indexical training. The shift to the symbolic reference strategy occurs when the chimp recognizes that the individual symbol/object references can be more economically learned and remembered when they are understood not as individual isolated associations, but as part of a systematically ordered representation of the object world encoded in the system of symbols or lexigrams.

In other words, the system of symbols is understood to function as a vastly more powerful reference system than the earlier and cumbersome process of acquiring reference to individual objects by personal acquaintance. Most of our social world is in fact dependent on symbolic categories, including institutions like money, marriage, and citizenship. As John Searle (1995: 70) explains, these institutions are inseparable from that of language, indeed “constituted” by language, which attributes to objects a *social function* that cannot be explained on the basis of either their physical or biological structure. For example, people can live together physically in a relationship that looks like marriage, but marriage alone institutes the requisite rights, duties, and obligations: these do not simply ensue from the fact of physical proximity but enter as a set of symbolic rules that we apply to it. Marriage may even exist without the partners cohabiting. All that is required is obedience to, or acceptance of, the symbolic bond.<sup>16</sup>

That much of our experience of the world is symbolic is not to deny that our cognitive makeup also inherits the bottom-up iconic and indexical representational strategies from our presymbolic past. Inevitably, a large part of our nervous system is dedicated to maintaining these earlier, more basic functions. But we also inherit a virtual or symbolic world—a cultural world—that involves not a bottom-up process, but a top-down one. We impose on the world categories and functions—symbolic meanings—that cannot be explained on the basis of the inherent physical and biological structure of the world and our bodies<sup>17</sup> but are instead *constituted* by the

16. I agree with Searle’s basic point that social institutions require a symbolizing function (i.e., language) because in order to represent duties, rights, and obligations, we have to go beyond the straightforward perception of objects. I think, however, that he underestimates the difficulty of explaining this shift from perception to symbolic representation. His formula “X counts as Y in context C” (Searle 1995: 55) is very far from being an anthropological hypothesis for the origin of representation precisely because it downplays the “evolutionary anomaly” of language: why should such a symbolizing function be selected for in the first place?

17. This opposition between top-down symbolic functions and bottom-up indexical functions is the equivalent of the old philosophical dichotomy between nature and culture. The latter is really a debate about the origin of meaning or, in evolutionary terms, the origin of intentional design or artifice in nature. See Deacon (1997: 89), who emphasizes the design-oriented or teleological structure of the symbolic system: “the shift from associative predictions to symbolic predictions is initially a change in mnemonic strategy, a recoding. It is a way of off-loading redundant details from working memory, by recognizing higher-order regularity in

process of symbolic reference.<sup>18</sup> It follows that a theory of symbolic culture must begin by explaining the origin of symbolic reference, which is to say, of language.

---

the mess of associations, a trick that can accomplish the same task without having to hold all the details in mind. Unfortunately, nature seldom offers such nice neat logical systems that can help organize our associations. *There are not many chances to use such strategies, so not much selection for this sort of process. We are forced to create artificial systems that have the appropriate properties*" [my emphasis]. Another way of putting Deacon's point is to say that symbols depend upon intentionality, which is evident in the artificiality of the symbolic reference system. Or in evolutionary terms: with the advent of the symbolic reference system, we get for the first time in evolutionary history the reality of intentional design. For example, archeologists who believe they have discovered physical evidence for symbolic abilities in early humans—e.g., highly regulated patterns or marks on fossilized bone—are responding positively (i.e., symbolically) to the following question: Were the marks produced by natural processes (e.g., by the claws and teeth of animal scavengers), or are they the result of intentional design, which is to say, of meaning? Note that this question is not even debated when the evidence in question is, for example, the sixteen-thousand-year-old cave paintings in Lascaux, because we assume that our own (symbolic) experience of the art is shared by that of the original artists: we have no difficulty in attributing intentional (aesthetic or ritual) design to the representation. The same interpretation applies, though admittedly less starkly, to the Oldowan stone choppers and flakes (about 1.9 million years old), which were presumably used for butchering meat by early hominids. The reason the distinction is less clear in the case of such Paleolithic tools is because the function of the tool can be predicted on the basis of the physical shape of the stone itself. But one cannot predict the function of the paintings from the physics and chemistry of the paint.

18. In suggesting that specifically anthropological categories are so constituted, I do not mean to imply that *all* ontological categories are symbolic. The failure to distinguish between specifically anthropological categories and more universal biological, chemical, and physical categories is a feature of cultural idealism. Cultural idealists hold that, because even scientific models of the world depend upon symbolic representation, we cannot assume a distinction between categories that preexist such representation and categories that are (in some sense) constituted by it. It would follow that science is every bit as constructive of the facts as religion or myth. For example, Fish (1980: 240) believes that there is no way to distinguish between "institutional facts" (e.g., the fact that certain bits of paper issued by the Federal Reserve count as money) and "brute facts" (e.g., the fact that those same bits of paper are cellulose fibers). The chemistry of paper is just another "set of discourse agreements . . . as to what can be stipulated as a fact" (ibid.: 242). Of course, in a trivial sense, Fish is right. Without symbolic representation, there could be no scientific explanation of the chemical structure of the world and no common meaning for words. But chemistry and literary criticism still differ in the level of reality (subatomic, atomic, molecular, cellular, psychological, social, cultural, etc.) that each discipline aims to interpret. It is not trivial to say that literature aims to constitute a community of readers by deferring reference to a nonsymbolic reality. But to apply the same functional explanation to chemistry is to impose on the ontological categories of atomic and molecular structure the causal model of anthropological explanation. As always, cultural idealism is typified by an excessive optimism concerning the constitutive power of language, an optimism shared by the cosmologies of mythical explanation. But to understand myth as a primitive form of natural science is simply a category error. What mythical explanation is ultimately concerned with is not the natural world "in itself" but the natural world as distilled through the lens of the always unstable ethical—i.e., inter-human—relationship. To the mythical mind, when disasters like earthquake and famine strike, it is never due to natural causes. It is rather the intentional design of some vengeful (divine or human) agent.

This conflict between top-down symbolic processes and bottom-up indexical and iconic ones points to a central difficulty for explanations of human cognition. The learning difficulties experienced by Austin and Sherman in their acquisition of a very rudimentary symbol system suggests not merely that chimpanzees are not naturally equipped to learn language, but also that our own cognitive evolution must have been confronted with a similar struggle—one furthermore in which we did not have the advantage of outside support from eager experimenters! It follows that cognitive science must sort out the tangled causal relations between an ontology based on the symbolic categories of language—and, I would add, culture—and those provided by neurobiology. And indeed, Deacon (1997: 298) emphasizes this predicament:

If we keep in mind that primate brains very much like our own have been around for tens of millions of years and that the mammal brain plan which our brain follows has probably been around for over 100 million years, it becomes evident that the logic of language is probably highly constrained to fit an ape brain logic. Though breaking up language analytically into such complementary domains as syntax and semantics, noun and verb, production and comprehension, can provide useful categories for the linguist, and breaking it up according to sensory and motor function seems easier from a global neuronal viewpoint, we should not expect that the brain's handling of language follows the logic of either of these categorical distinctions. The patterns we observe probably reflect, in a very indirect sort of way, the processing problems produced by mapping a symbolic reference system encoded in a serially presented modality onto the processing logic of an ape brain.

The consequences of this dichotomy between basic sensorimotor processes and higher-order symbolic ones are significant. Failure to recognize it leads to a wrongheaded scientism, on the one hand, or to airy spirit mongering, on the other. Certainly, one can sympathize with Turner when he argues for a genuine *rapprochement* between cognitive science and literary studies. And yet his own attempt to create such a cognitive poetics reveals why such attempts will remain at best incomplete, and at worst simply misguided, if they do not absorb the central lesson that symbolic processes cannot be sufficiently explained on the basis of their underlying neural-cognitive processes. What is needed is not just a poetics inflected by cognitive science but a cognitive anthropology.

#### 4. Cognitive Anthropology

In this section, I will explain why the cognitive model of representation, as elaborated by Turner (1996) and by Lakoff and Johnson (1980), is inade-

quate by juxtaposing it with Deacon's idea of symbolic reference. More precisely, I will argue that it is the *evolutionary anomaly* of symbolic reference that the cognitive model overlooks in its account of representation. By assuming that such symbolic categories as narrative<sup>19</sup> and metaphor are selected for, just like any other biologically useful sensorimotor function, the cognitivists confuse specifically anthropological with more universal biological categories. In their enthusiasm for cognitive science, they see the origin of institutions like literature and art in the deep structures of the hominid mind. But this move from the institutional structure of social reality to the innate mechanisms of the mind seems premature at best.

The notion that the cognitive and neurological sciences can provide a sufficient account of the origin of symbolic processes depends on the mistaken assumption that symbolic reference is a genetically assimilable trait.<sup>20</sup> Whereas indexical signals, such as vervet monkey calls, possess the long-term referential stability that allows for genetic assimilation, so that over the course of many generations these calls become innate, symbolic reference is defined by the fundamentally open-ended nature of its reference. What characterizes the symbol is its relationship to other symbols. The sign's symbolic reference—its meaning—is thus essentially “parabolic” (see Turner 1996: 7), dependent upon the relations or “projections” within the entire network of symbols. When a new meaning is introduced into the system,

19. As Genette (1980: 26–27) has shown, the analysis of narrative assumes a symbolic relationship between signs and events. But beyond assuming the formal necessity of this relationship, narratology has given little attention to what motivates it. My assumption is that the significance of this relationship is most fruitfully understood by way of a hypothesis of their coeval origin.

20. In terms of genetic assimilation or Baldwinian evolution, for example, a behavioral pattern of recurring migration to cooler climates can select for individuals with warmer coats should the cooler region become increasingly important for survival; or more dramatically, invariant features among sign, neural processing, and environment can lead to the hard-wiring of vervet monkey alarm calls. Applied to the issue of language origin, this would mean that the universal features of language postulated by cognitive scientists like Pinker (1994) as candidates for adaptive success must be susceptible to Baldwinian evolution. But Deacon (1997: 330) shows that there is in fact no correlation between the deep grammatical structure of language and the structure of neural processing and concludes that the genetic assimilation of grammar by Baldwinian evolution is biologically implausible. Searle (1992) reaches a similar conclusion when he argues that the postulation of deep unconscious rules to explain language acquisition confuses the distinction between causal and functional explanation. In seeking a causal explanation, there is no reason why we should assume that the underlying unconscious neural mechanisms should mirror the grammatical patterns of language. What Searle objects to is the move from a functional explanation, based on observable grammatical patterns, to a causal explanation that must reproduce those patterns at some *deep unconscious* level. But postulating deep unconscious rules of universal grammar adds no “further predictive or explanatory power,” and so the explanation is not in fact “causally efficacious” (Searle 1992: 244).

it is always on the basis of the existing network of relations, from which the new meaning is “projected.” But the socially constructed nature of the symbol system, that is, the fact that for a sign to be understood its meaning must be accepted by other symbol users, prohibits the genetic assimilation of symbolic reference. Meaning is essentially unstable because it depends upon a symbol-using community. The acquisition of symbolic reference is thus not an inevitable outcome of the genotype. This, more than anything else, points to the central evolutionary problem of language origin: what “makes symbolic associations so difficult to learn, also makes them *impossible to assimilate genetically*” (Deacon 1997: 332). Or, as the cognitive psychologist Merlin Donald (1998: 50) says: “Single brains simply do not invent languages or symbols, and in isolation do not appear to have the capacity to do so.”

In anthropological terms, this impossibility of genetic assimilation rules out a smooth and gradual transition from iconic and indexical to symbolic reference. But this is exactly what Turner assumes in his study of the literary mind. By reducing symbolic processes to cognitive processes and—ultimately—to brain processes, Turner manages to eliminate from his ontology all anthropology. In effect, Turner argues that symbolic thought is “nothing but” patterns of mapping in the brain.<sup>21</sup> He believes that the brain represents concepts fuzzily, that is, in more-or-less terms depending on the number and strength of the neuronal connections that are stimulated during consciousness. Instead of our concepts having clear and distinct boundaries,<sup>22</sup> they constantly grade into one another. Thus, just as the color red shades gradually into orange, so that we see not one red but a gradation of reds as we move along the color spectrum, so too our concepts merge together to create new concepts or “conceptual blends.”

But in this Turner has confused the symbolic map with the neurological territory. No doubt, as Deacon’s (1997: esp. chap. 10) research demonstrates, it is implausible to expect a neat one-to-one mapping between the macrostructures of linguistic function (e.g., semantics and syntax) and the microstructures of neural processing.<sup>23</sup> But nor should we therefore conclude that

21. This is particularly evident in *Reading Minds*, where Turner (1991: 46) claims that certain patterns in the brain are “inherently meaningful.” But this is not to speak of “human meaning” at all.

22. Cf.: “A concept does not have hard edges. If we think of a concept as an activated set of links in a pattern, then different links in the pattern will have different degrees of strength, and there will be no clear boundary to how strong any link must be to qualify as belonging to the pattern. Concepts fade out at their boundaries, as opposed to stopping abruptly” (Turner 1991: 45).

23. Cf. esp. Deacon 1997: 288: “Once we abandon the idealization that language is plugged into the brain in modules, and recognize it as merely a new use of existing structures, there is

the anthropological macrostructures of symbolic reference reduce to the cognitive microstructures of neural processing. It may well be that, in terms of neuronal processes, meaning is far from static, and no doubt word meaning is differently represented in different brains, as Deacon's research suggests. But this does not mean that, at the *symbolic level*, our concepts lack categorical boundaries. It all depends on how one understands the origin and function of symbolic categories.

As I have argued, symbolic categories have a specifically anthropological origin and function, irreducible to the basic categories presupposed by a cognitive science grounded in the biological functioning of the brain. From the perspective of such a cognitive science, categorization can be explained in terms of simple iconic and indexical recognition processes, which require no more than a functioning nervous system. Cognitively speaking, the capacity to differentiate perceptual experiences at the iconic level implies the ability to interpret perceptual experience indexically, categorizing sensory experience in such a manner that new experiences may be anticipated and thus acted upon. In this sense, icons function not only as icons of themselves but also as indices of more complex iconically differentiated patterns. In recognizing smoke, for example, one refers this experience not merely to past icons of smoke but to a whole sequence of iconic representations, including other icons for smoke and for fire; it is the memory of this overall pattern that enables an instance of smoke to function as an index for fire. But understanding the origin and function of iconic and indexical referential processes will not help us explain the origin and function of symbolic reference. This is the error Turner (1996: esp. chap. 2) makes in his attempt to explain the origin of "human meaning" in terms of the structure and function of the brain. The meaning of the symbol is a social or collective imposition that is not based on the intrinsic features of the individual brain, the sign, or the object signified.

It is important to recognize this distinction, because the whole question of the origin and function of social (i.e., anthropological) categories depends upon it. If we understand social categories to originate in the individual mind, in the self's internal representation of the external environment, then we may assume that social categories are indeed merely "abstract" or "metaphoric" projections of what the individual can perceive independently of what anybody else can perceive (see Lakoff and Johnson 1980: 59; Turner 1996: 18). The problem then becomes how *my* representation of

---

no reason to expect that language functions should map in any direct way onto the structural-functional divisions of cortex. It is far more reasonable to expect language processes to be broken up into subfunctions that have more to do with neural logic than with linguistic logic."

a particular category can correspond to *your* representation of it. Cognitive science tends to assume that this problem is explained by the species-wide uniformity of ontogenetic development: the interaction of genetic and environmental factors so constrain us all that there can be no significant difference between the kinds of categories each individual is able to perceive and thus represent symbolically. I do not doubt that, at the perceptual level, we do indeed make more or less the same distinctions and categorizations. What I disagree with is the implication that symbolic representation emerges naturally or gradually from perceptual experience. The imposition or “projection” of symbolic meaning onto directly perceivable objects or referents is a socially mediated activity, and as such it cannot be explained purely on the basis of individual cognitive processes (Turner 1996: 7, 17). Without a model for the collective or social function of symbolic reference, we cannot understand the origin and evolution of human cognition.

In fact, social anthropologists have long recognized that social categories, such as kinship relations, function independently of the real biological relationships on which they are imposed.<sup>24</sup> Nonhuman primates can doubtless categorize immediate kin, but the imposition of a kinship system requires more than higher perceptual abilities and group living. Without language (or symbolic reference), it would be impossible to turn the ad hoc representations of the perceptual system into a rule-governed system. As many anthropologists have noted (see, e.g., White 1959, Fox 1980, Knight 1991), the function of kinship systems is to *constitute* or *define* social categories in order to regulate behavior in a context that transcends mere indexical or perceptual proximity. A dominant male baboon can maintain his monopoly over the females only by physically fending off other males, including his own male offspring. In the human case, the institution of the incest taboo and—derivatively—of out-group marriage accomplishes the same function, but only given the capacity for symbolic representation. Thus, in order to explain the incest taboo and the system of differences on which the kinship system rests, we must first explain the origin of language and the symbolic differences thus generated: between wife and mother or husband and son, for example.<sup>25</sup> Failure to recognize this boundary can be disastrous for the individual, as Oedipus’s tragedy of unknowing transgression illustrates. Knowing how this difference is instantiated in the brain does not explain

24. For example, Fox 1983: 34: “A consanguine is someone who is defined *by the society* as a consanguine, and ‘blood’ relationship in a genetic sense has not necessarily anything to do with it, although on the whole these tend to coincide in most societies in the world.”

25. White (1959: 82–83) recognizes the inseparability between the origin of language and the origin of kinship systems but tends to regard the former as a wholly unproblematic evolutionary occurrence.

its social function, which, I would argue, stands at the origin of the distinction.<sup>26</sup> I will return to this point in the final part of my argument.

Why does Turner make this error of conflating the social function of representation with the tendency to generalize on the basis of particular sensorimotor categories? Why does he reduce symbolic to cognitive and ultimately to neurological categories? Turner commits a variety of what Searle (1992: 36), in his critique of materialism, calls the identity error, namely, the belief that intentional (i.e., symbolically represented) states are *nothing but* neurobiological events. In order to understand a particular word, it is of course necessary for its meaning to be represented in, and thus caused by, my brain. But the mere fact that neuronal processes cause our conscious experience of symbolic meaning does not mean that our experience of symbolic meaning is *identical* to these neural processes. Another way of putting this is to say that meaning is not reducible to brain processes because meaning—symbolic reference—is a collective or social reality irreducible to any individual brain state.<sup>27</sup> In arguing for such a reduction, Turner (1996: 7) leaves out precisely what makes metaphoric projection—“parable”—possible: the mediation of individual experience by the collectively imposed and collectively shared sign. This is also why Deacon can declare that symbolic reference is genetically *inassimilable*. In the fundamental case of sacred or totemic prohibitions, for example, categorical interdiction of an indexically perceivable object is a symbolic imposition possible only because such an interdiction is accepted by each member of the symbol-using community. How this symbolic association is made in the

26. At this point, it may be objected that I am simply misreading the cognitivist project, which aims to attack classical metaphysics and its rigid adherence to mind/body dualism. How can I deny that symbolic thought must make use of more biologically primitive cognitive functions? In developing a theory of representation, surely we must first study how categories are formed at this basic cognitive level? For example, on the basis of Eleanor Rosch's work, Lakoff (1987) argues that our concepts exhibit “prototype” effects. Thus we tend to think of sparrows as more representative or more prototypical of the category *bird* than penguins, which suggests that “prototypes act as *cognitive reference points*” from which inferences can be drawn (ibid.: 45). However, my point is not to deny that we make use of prototypes in drawing inferences, but that we cannot take such inferential processes as a sufficient explanation of symbolic reference.

27. I do not wish to imply that cognitivists simply ignore cultural influences on the individual. On the contrary, they frequently acknowledge that all kinds of external pressures, from material bodily needs to sociocultural conventions, shape cognition. But never is there any reflection on the ontological difference between perceptual and cultural representations. Instead of being explained, this difference is assumed, for example, in Lakoff and Johnson's (1980: 178) distinction between direct experience and indirect experience or Turner's (1996: 4–5) distinction between story and parable. I think the reason for this omission is obvious: cognitivists interpret the emergence of culture as a gradual outgrowth from more basic sensorimotor functions. For them, ultimately, culture must be reducible to an ontology of the brain, the only other alternative being a return to old-fashioned mind/body dualism.



brain is secondary to the public or *scenic* nature of the interdiction itself. By tracing the origin of the symbol back to the brain, and not to such a collective scene, Turner foregoes the very possibility of explaining the uniqueness of human cognition. Yet he continues to assert, quite paradoxically, that his account shows what is unique about human thought.<sup>28</sup>

How are we to explain this paradoxical situation in which a putatively “anthropological” understanding of human cognition succeeds only by eliminating all trace of anthropology from its basic ontology? To answer this question, we need to examine a parallel “hopeful monster”<sup>29</sup> strategy from linguistics. Linguists (see, e.g., Chomsky 1972, Pinker 1994, Bickerton 1995) have long been impressed by the universality of certain features of language,<sup>30</sup> which allegedly correspond to universal features of the brain. The strong innatist hypothesis for a language organ or universal grammar is a consequence of this kind of reasoning. Turner (1996: chap. 8) does not believe in a language organ, but his argument for the innate existence of his basic concepts of narrative, projection, and parable is motivated by a similar projection of symbolic structures onto brain structures. In this sense, he is a victim of his own metaphor for human cognition, namely, “parable”:

28. For example, after asserting that “story, projection, and parable . . . are the root of human thought,” Turner (1996: 12–13) gives the following definition of story: “The basic stories we know best are small stories of events in space: The wind blows a cloud through the sky, a child throws a rock, a mother pours milk into a glass, a whale swims through the water.” To the objection that these are not stories at all but simply evidence of basic sensorimotor abilities, Turner (*ibid.*: 13) responds that, given the fact that “five billion different human beings all recognize and execute” these “small spatial stories,” “we have a hard time imagining that the capacity can be interesting.” But, Turner (*ibid.*) continues, we should not be misled by appearances, because this is in fact “the chief puzzle of cognitive science” and “if you do not have this capacity [i.e., for producing small spatial stories], you do not have a human mind.” But one cannot dismiss the distinction between the commonsense and the “scientific” notion of story simply by stating that science inquires into “boring” (*ibid.*), unnoticeable things (like perception) rather than the exciting things (like Achilles’s wrath or Goldilocks’s encounter with three bears). In grounding the primacy of his cognitivist notion of “story” in its anthropological universality, Turner fails to realize that his idea of story is biologically universal as well. The ability to remember sequences of perceivable events, like clouds blowing through the sky or milk pouring into a glass, is hardly unique to human beings and certainly very far from our normal understanding of what a story is. His claim that such capacities are indispensable for having a human mind therefore seems overstated at best, and simply mistaken at worst, since these are the conditions for having any conscious mind at all, whether human or animal.

29. The term *hopeful monster* is due to Richard Goldschmidt (see Dennett 1995: 288); it refers to the idea that evolution can throw up dramatic macromutations leading to radically new forms and lines of descent for micromutation to work on (see Goldschmidt 1952 for a review of the idea). The most famous “hopeful monster” theory of human language is Noam Chomsky’s idea of an innate universal grammar (see Deacon 1997: 35).

30. For the counterargument to the innateness hypothesis, as well as for a useful review of the main issues, see Sampson 1997.

as Turner (*ibid.*: 7) explains, it comes from the Greek, *paraballein*, meaning “the tossing or projection of one thing alongside another.” It is not that Turner is wrong about the dominance of visual and manual metaphors in our thinking. He is just wrong to see this dominance as an explanation for the origin and function of symbolic reference.

It is in fact quite unsurprising that preexisting sensorimotor patterns should be co-opted by symbolic functions. Metaphor is essentially a process of recognizing the similarity between different domains of *symbolically mediated* experience. For example, in perceiving the similarity between a laser beam and a slap shot by Wayne Gretzky, I might say, “Gretzky scored with a laser in the final minute.” At the linguistic level, metaphor occurs as a substitution of the word *slap shot* with *laser*. Note that I cannot substitute just any word. I cannot say, “He scored with a house in the final minute.” The key to metaphor, as David Lodge (1977: 75) notes, is the bringing together of similarities with the “awareness of *difference*.” At the presymbolic level, we can of course make associations on the basis of perceptual similarities. As I have already explained, such associations are based on indexical and iconic reference, and they enhance the predictive powers of all cognitively aware organisms.

With the origin of symbolic reference, however, the possibilities for the association of perceived similarities is vastly increased to include our generalizing from totally distinct perceptual experiences. This form of generalization is quite distinct from stimulus generalization and indexical association. As Deacon (1997: 88) argues, it is really a form of “logical or categorical generalization” because what mediates our sense of the similarity between two disparate experiences is not a perceptual association but a symbolic one between different paradigmatic (i.e., semantic) or syntagmatic (i.e., syntactic) possibilities. In metaphor, the association is paradigmatic; a word is substituted for another with similar semantic features. For example, if I say, “His lecture ignited the room,” I am projecting the experience of fire onto a completely different experience, namely, a thrilling lecture. It is tempting to see this association as an example of stimulus generalization, based on the similarity between our experience of fire and of a stimulating lecture. But what fires and lectures share in common is not any physical property, but the fact that they can be referred to *symbolically*. It is the massive leap in associative possibilities enabled by symbolic reference that allows us to draw a metaphoric association between vastly different perceptual experiences.

Deacon (1997: 303–6) notes that metaphor can be traced to the posterior cortical regions of the brain. This should not surprise us, he suggests, because “associations cued by attention to common perceptual features are analogous to perceptual recognition processes and so should recruit the

function of the corresponding posterior cortical regions” (ibid.: 306). In other words, metaphor is a species of symbolic reference that relies heavily on similarity. In constructing a metaphor, we recognize a semantic or paradigmatic similarity between two different words. At the symbolic level, the capacity for recognizing common features between otherwise disparate perceptions is vastly increased, because we are now predisposed to interpret the world in terms of symbolic, rather than iconic and indexical, relationships. When Hamlet tells Polonius that a cloud looks like a camel, or a weasel, or a whale, he is putting iconic recognition processes to a highly abstracted symbolic use. Only humans could make such abstract use of iconic processes because only humans have evolved the ability to think symbolically.

Language seizes on the metaphoric possibilities offered by our presymbolic visual and manipulative biases. But from the latter’s ubiquity, we should not conclude that they are themselves the source of our capacity to think symbolically. What remains unexplained in Turner’s account is why prelinguistic cognitive modes need to be represented symbolically.

It should be evident that this is the same criticism as I made earlier of Lakoff and Johnson (see section 2). There I suggested that the problem with their account is that they pursue the origin of metaphor solely from within the explanatory framework of ontogenetic development: their discovery of systematic biases toward spatial and bodily metaphor in linguistic usage is understood as reflecting deep biases in psychological and cognitive development. But the developmental perspective cannot address the question of the origin of metaphor. All empirical study of our metaphoric biases already assumes the process of metaphoric substitution, which is to say, ultimately the origin of language as symbolic reference.<sup>31</sup> And the simplest proof of this need for a hypothesis of anthropological origin is that language is not invented by children but inherited from adult language users.

31. As Miers (1992: 969) notes, Turner (1991) cannot appeal to the sheer weight of his empirical data because his interpretation of this data already implies a theory of representation—in Turner’s case, a theory of meaning grounded in a “natural link between language and the body” (Miers 1992: 967). Miers’s (ibid.: 967) main point is that the “cognitivism” espoused by the likes of Turner and Lakoff mistakenly interprets “the presentation of embodiment as if it were not already a re-presentation.” I agree with this criticism, but I think that Miers too ultimately underestimates the anthropological nature of the problem of representation: he suggests that the discovery of its “neural recipe,” as highly distributed processes in the brain, ends the “dialectic adventure” begun by Hegel and continued by Derrida and launches a new theory of representation grounded in a genuine cognitive science, which Miers (ibid.: 954) is always anxious to distinguish from its “reactionary” use by Lakoff and Turner. To say the least, I am skeptical of the possibility that cognitive science, whether reactionary or revolutionary, can provide us with an understanding of representation superior to that of either the dialectic or the analytic tradition without at once also taking into consideration the *specifically anthropological context of this problem*.

The same critique applies to Turner. From the evolutionary and anthropological perspective, his suggestion that story, projection, and parable provide “a cognitive basis from which language can originate” (Turner 1996: 168) merely begs the question. We cannot invoke concepts like story, projection, and parable, because these concepts are precisely what needs to be explained since they are all instances of symbolic processes. Turner posits these symbolic processes as ontologically independent of symbolic reference and then derives the fact of symbolic reference from them. This is exactly what Lakoff and Johnson did to explain the origin of metaphor.<sup>32</sup> In each of these accounts, symbolic reference is assumed but not explained.

As I have already pointed out, Turner makes strong biological claims for his three categories of narrative, projection, and parable. These concepts, he argues, are fundamental in the sense that they evolve biologically as mental features of the brain. Hence his claim that they are selected for, just like color vision and object recognition (see Turner 1996: 5, 14–17). I have argued, on the other hand, that it is impossible for narrative *as a symbolic process* to be genetically assimilated along Baldwinian lines.

But what about Turner’s more basic claim that narrative is simply an adaptive feature of perceptual and motor processing? Isn’t this a feature that can be genetically assimilated? My answer is that these more basic sensorimotor functions are certainly part of our biological inheritance, but then we are not talking about human narrative anymore. Instead, we are talking about basic sensorimotor functions. These are interesting in their own right, but they must not be confused with narrative or metaphor or any other symbolic process.

The distressing thing about Turner’s analysis is that he oscillates, apparently without realizing it, between two versions of narrative that are entirely incompatible.<sup>33</sup> On the one hand, narrative is uniquely human; on the

32. Perceptive readers will note that I have omitted an intermediary step, in which Lakoff and Johnson first separate linguistic metaphor from conceptual metaphor. Conceptual metaphor is then grounded in nonmetaphoric concepts that, the authors claim, emerge from “direct physical experience” (Lakoff and Johnson 1980: 57, 58). Thus, Lakoff and Johnson’s schema goes as follows: direct experience → emergent concept → conceptual metaphor → language. It is revealing, however, that when the authors refer to direct physical experience, they do so only to embed the function of metaphor *within the very idea of experience*. Thus they (ibid.: 56, 58) speak of experience as generating “emergent concepts” and “emergent metaphors,” as if language could be expected to emerge miraculously from basic sensory and motor functions. This is an instance of rhetoric reproducing the paradox of the origin, which no doubt a Derridean reading of Lakoff and Johnson would be quick to point out. The deconstructionists have a point, but I am far more optimistic than they about the possibility of a constructive solution to the paradox of symbolic reference.

33. Herman (1999: 26) also remarks on Turner’s careless definition of narrative when he notes that narrative always involves “the remarkable and the tellable” and not “the stereotypi-

other, it is nothing but basic perceptual and motor patterning. One of his examples of a story is throwing a rock (Turner 1996: 17). The brain's ability to process this "narrative" enables us to duck when we see someone aim to throw an object at us, because we interpret the raised arm as an *index* of the entire sequence of throwing. It is of course easy to see how such behavior is adaptive and so could lead to an innate competence for evasive action (e.g., blinking, flinching, and so on). But here is the conundrum: we duck when we see a raised arm, but so does a monkey or a bird or a dog. This is not a form of projection in the specific "literary" sense that Turner discusses in the rest of his book. It is a form of pattern completion or inference that, as Deacon explains, requires no higher complexity than indexical reference, and indexical reference is supported by all nervous systems, not just our own.

Turner evidently wants to have his cake and eat it too. If his concepts of narrative, projection, and parable are truly such elementary cognitive processes, like color vision and object recognition, then we share them with many other species. But if that is the case, then narrative—to take only his most basic concept—is not uniquely human, and his claim that cognitive science stands at the center of the human sciences must be taken with a grain of salt. In his optimism for the cognitive paradigm, Turner paradoxically erases from his discussion precisely what is human about narrative.

What, then, is to be learned from Turner's study of the literary mind? Turner's most suggestive and illuminating discussions take place when he analyzes our everyday cultural metaphors and blends. These analyses stand on their own, independently of the wider and more dubious ontological claims he makes about the origin of our literary processes in general. As I have sought to emphasize, his claim that narrative, projection, and parable are independent of language—of symbolic thinking—is deeply mistaken. Given his schema, the really crucial element is projection. But what Turner means by this concept is explained, far more adequately, I think, by Deacon's theory of symbolic reference. Projection is the symbolic capacity to think in terms of symbol/symbol relationships independently of their symbol/object associations, on which, of course, together with their more basic iconic relations, they nonetheless ultimately depend. Naturally, we can project concepts independently of language; we do not have to speak out loud or even internally in order to make symbolic associations. But we cannot make these virtual associations independently of symbolic reference. When I use a map to find my way from King's Cross to Covent

---

cal and expected." Herman, however, does not elaborate on why symbolic reference always involves "the remarkable" instead of the predictable.

Garden, I am making symbolic associations, without speaking any natural language. Turner's idea of the blended space (see Turner 1996; Turner and Fauconnier 1999), in which already existing concepts are combined to generate new meanings, is thus itself a metaphor for this symbolic world—which is to say, ultimately for the origin of language in the broad anthropological sense. Symbols are indeed “nothing but” metaphors, if by metaphor we mean simply the fact that words are necessarily implicated in a network of relations with other words. In the final analysis, word-meaning is nothing but projection, because in order to refer symbolically, a word must be projected onto a separate mental space where it can be thought independently of the object-world to which it nonetheless ultimately refers.

We can better appreciate Turner's analyses of language once we get past the cognitive science to the anthropological reality to which his analyses ultimately refer. The moment of specifically human cognition occurs when a particular image or story is mapped or projected from one context to a different context, as when we compare the passage of time to a flowing river, or a courageous man to a lion. As Turner implies, our concepts of the world are essentially poetic or creative because the construction of these concepts is mediated by a symbolic system that projects from a vast storehouse of symbolic associations. But the symbolic system itself is not simply an outgrowth of our more basic sensory and motor systems. Symbolic reference is a projection produced in the first place collectively; only if a group of symbol users—the minimal linguistic community—accepts the reference of the sign can symbolic projection exist. Ritual is designed to maintain the public status of individual (sacred) signs. The inflexibility of ritual is a recognition of the fact that these signs are fragile precisely because they do not possess the concrete predictability informing our direct, nonsymbolic sensorimotor experience of the world. Religious icons and stereotypically performed rituals are concrete historical indices of a world that is, in the final analysis, only available symbolically.

### **5. The Anthropological Origin of Symbolic Reference**

Thus far, I have argued that symbolic reference is the key to understanding the difference of human from animal cognition. Theories that explain symbolic reference as evolving gradually from more basic cognitive capacities, such as perceptual association and categorization, are doomed to failure, because these basic capacities are all explainable on the basis of a simpler model of representation—indexical reference—than the symbolic. From the originary perspective, the transition from index to symbol cannot be regarded as an inevitable or highly predictable consequence of ontogenetic

development, for the simple reason that at the origin no such model of ontogeny can be presumed. It is rather *the origin of symbolic reference itself* that acts as the major selection pressure on the evolution of the genotype and its phenotypic effects in ontogeny. Only after a long period of brain-language coevolution can we assume the stabilization or “adaptation” of the human genotype for symbolic functions. This significantly changes the way we pose the question. Instead of asking, “What presymbolic cognitive mechanisms are adapted for symbolic functions?” we must ask, “What indispensable anthropological function does symbolic reference perform that has pushed it to its current status as a universal feature of our species?”

Let me emphasize this last point. I am saying that specifically human cognitive functions, like metaphor and narrative, are a *consequence* of the capacity for symbolic reference. Simpler cognitive functions, like perceptual inference and basic categorization, must be understood as co-opted by this qualitatively different, symbolizing function. But if this is the case, how do we explain the origin of the symbolizing function? It appears that the theory must be based on a paradox: I must be claiming that symbolic reference is only explainable in terms of itself.

This is indeed my claim, and this is also the reason why we require an anthropological hypothesis for the origin of symbolic representation, instead of locating it in the underlying cognitive and neural mechanisms. As Deacon (1997: 447) argues, symbolic reference is not something intrinsic to the brain but a peculiarity of the way humans interpret their relationships, both to one another and to the world in general (one might view the former as the minimal anthropological or ethical relation and the latter as the maximal cosmological or sacred relation). It follows that a *minimally sufficient* explanation of symbolic reference must be a functional one. And the minimal function of symbolic reference is, quite simply, symbolizing.

A less-parsimonious way of putting this is to say that symbolic reference assumes intentionality: the capacity to attribute intentional mental states not only to other similar beings (e.g., other humans) but also to objects (e.g., words on the page, totems, sacred animals and objects, etc.). But in order to attribute intentionality, you need to have something to think with, that is, signs or tokens that have meaning or symbolic value attributed to them. In other words, you need to have language, a system of signs whose relationships are symbolically defined, rather than learned by indexical association. It follows that ritual (i.e., the interdiction of worldly objects through their sacralization) assumes a symbolizing function. Ritual is an elementary or primitive form of language.

Having intentional mental states entails in turn the potential to influence evolution in an unprecedented manner. Just as a dog breeder inten-

tionally selects for certain characteristics that will enhance the usefulness of dogs for particular tasks (e.g., hunting or sheepherding or sled pulling), so symbolic reference selects for symbolizing abilities and imposes its own selection pressures on the brain. It is for this reason that the mechanism of natural selection (i.e., the replication of “selfish” genes [Dawkins 1989 (1976)]) cannot explain the emergence of symbolic representation. The symbolic function is instead the originating factor in a *coevolutionary* process involving human culture, on the one hand, and human biology, on the other. Within this coevolutionary process, culture can only be explained in anthropological-functional terms, geared to the use to which symbolic reference strategies are put.

Note that I am not denying that symbolic reference requires more elementary cognitive and neurological processes, which are themselves explainable as biologically evolved adaptations. Rather, I am saying that the origin of symbolic reference can only be grasped in terms of a hypothesis of its underlying function, which alone can explain how the adaptation for language (the “language instinct” [Pinker 1994]) got to be there in the first place. Obviously, an increasing reliance on the symbolizing function exerts a strong selection pressure on brains to *adapt to* this function. But the adaptation itself does not explain the function. Furthermore, once this coevolutionary process has been driven to stabilization—to the point where further modifications at the genetic level are no longer optimal—we are still left with the need to explain the originating process that led to this genetic adaptation. Genetic adaptation for symbolic abilities does not lead to a stasis in the overall system. On the contrary, symbolic reference contains within itself its own generative dynamic. Thus, it not only evolves in parallel to its biological and neurological substrate, it also far outpaces it in its representational and, consequently, in its adaptive capacities. With the origin of symbolic reference, there emerges the possibility for a new (i.e., nongenetic) kind of information transmission: the transmission of culture.

To sum up, cultural evolution is the consequence of intentional actions. The very first symbolic act must have been intentionally performed: it had a symbolic function and was performed only because it had that function. A hypothesis for the origin of symbolic representation must therefore begin with an account of its function. Since this function is specific to one species—*Homo sapiens*—it must also be uniquely anthropological. Thus, in order to explain the transition from index to symbol, we need a hypothesis for the anthropological function of symbolic representation.

Deacon himself—perhaps surprisingly for an empirical scientist—seems quite happy to admit the need for such a speculative “originary” hypothesis. In the third and final part of his wide-ranging argument, he answers



the functional question by hypothesizing that the symbolic sign originates in response to the peculiar ecological predicament of our meat-eating and group-living ancestors. In particular, the conflict between increasing group size, required to exploit the hunting possibilities of the open savanna, and the need for groups of male hunters to ensure their paternal certainty, so that they will not support someone else's offspring, is solved by the introduction of the symbolic bond of marriage (Deacon 1997: 393–401). The crux of his argument is that the symbolic sign allows a relationship to be constructed that can function independently of the physical proximity needed for indexical sign/object associations. The symbolic bond of marriage is in effect a promise on the part of the individual vis-à-vis the entire group to sacrifice the possibility of unlimited promiscuity (which requires sheer physical dominance over competitors) for the more reciprocal or “moral” situation of a socially sanctioned marriage.

This kind of “social contract” theory is of course not new. Beginning with Thomas Hobbes (1947 [1651]), philosophers have frequently speculated on the origin of social obligations from the “state of nature.” What distinguishes Deacon's account from thinkers like Hobbes and Jean-Jacques Rousseau (1994 [1755]), however, is his attempt to situate the origin of the social contract in that of language. The potential for symbolic representation is immanent in the mutually contradictory pressures of increasing group size and regulating the conflict between males that such an increase inevitably leads to. We may remain skeptical of such a scenario (see, for example, Ruse 1998), but we must remember that this is all that can be expected of any evolutionary explanation, always formulated after the fact.

In the comparative context of evolutionary biology, however, the case of the origin of symbolic functions is unique, because the origin can only be approached from the “emic” perspective of symbol users interpreting the behavior of other (hypothetical) symbol users. There is no *outside* comparative perspective from which to measure the functional and evolutionary success of symbols. Comparison of similar traits across different species is of course the basis of evolutionary biology, but such a comparative method is insufficient for a hypothesis of the anthropological function of symbolic reference. Instead, all confirmatory evidence for it must be *anthropological*, that is, intrinsic to the cultural and historical record of humans since the first moment of symbol use. Competing hypotheses can only be accepted or rejected on the basis of their power to explain this cultural record more or less satisfactorily.

This is in fact no more than has traditionally been conceived for humanistic study, which begins by interpreting texts rather than the objects of unmediated experience. But in contemporary cultural anthropology, as in

the humanities in general, the emic status of symbolic reference (i.e., the fact that any theory of language origin must inevitably assume the very categories it seeks to explain) has tended to produce the pessimistic view that the question of language origin is unanswerable (as in Derrida 1974). Paradoxically, this is the same conclusion reached by impeccably empirical scientists, though for radically different reasons. The scientists have no difficulty with representation; it is rather the excessive burden posed by hypothesizing an unfalsifiable event or moment of origin that offends their idea of a legitimate scientific enterprise.

To the scientists, one need only point out that speculative hypotheses are integral to any advance in scientific understanding. As Marvin Harris (2001 [1968]: 286) has pointed out, the insistence on the primacy of data collection, to the neglect of theoretical speculation, does not lead to scientific advancement. To the pessimists in contemporary cultural theory, the idea of a positive solution to the problem of representation is less likely to be welcome, if only because global theories of cultural origin and evolution tend to be associated with the Darwinian excesses of nineteenth-century anthropology, which is in turn explained as a consequence of nationalism and imperialism.<sup>34</sup>

I think this dual dismissal of the origin of human linguistic and cultural capacities exposes a fundamental contradiction in the nature of anthropological research that nonetheless points us in the direction of a synthesis. This synthesis, however, is unlikely to come from the sciences, though Deacon clearly comes very close to extending a hand to humanistic inquiry in his recognition that the human brain is selected for symbolic reference, rather than the other way around.<sup>35</sup> His coevolutionary argument puts lan-

34. In his masterly study of "the rise of anthropology," Harris (2001 [1968]: 250, 287) points his finger at Franz Boas, whose "historical particularism" was "inductive to the point of self-destruction." Boas, Harris argues, was reacting to the evolutionism bequeathed to anthropology by the great nineteenth-century anthropologists: Herbert Spencer, Edward Tyler, and Lewis Henry Morgan. White (1959: 70–72) similarly notes that "the schools of Boas, Radcliffe-Brown, and Malinowski" have all rejected the problem of the origin of culture and deride evolutionary theories of culture as "conjecture," "unverifiable," or "sheer speculation," but they fail to recognize that all scientific hypotheses are "speculative" in the sense that they are "built up with inferences." In anthropology, the polemic between evolutionism and historical particularism continues to this day. To scholars in the humanities, this polemic may seem somewhat passé. Are we not all historical particularists in the Boasian sense? The very notion of literature departments divided along lines of national cultures implies the historical specificity—rather than the anthropological universality—of cultural explanation. Some rear-guard actions continue to be fought in the name of liberal humanism: see for example Good's (2001) critique of the rise of "theory" in literature departments. But despite Good's misgivings, the message underlying his analysis is clear: historical particularism has won the day.

35. The difference goes to the heart of the language-instinct debate. Is language innate, or

guage and culture at the forefront of biological anthropology. What remains to be done is to understand the origin of language and culture in terms of a minimal hypothesis that can at the same time form the basis for the interpretation of historically significant works of culture and literature.

The latter is of course the bread and butter of students in the humanities. And the humanities, in contrast to the sciences, have always been receptive to empirically unfalsifiable but anthropologically suggestive ideas (witness the dominance of Freud). The success of deconstruction lies largely in its idea of representation as a deferral of the immediate possibility of falsification; symbolic reference succeeds only by deferring its relationship to the world in favor of its relationship to other signs. This is an intuitively attractive idea to the literary critic, because it confirms our experience of reading fictional texts. But the notion of deferral, if it is to be more than the source of a philosophical critique of logocentrism, must be grounded in the anthropological reality that is our capacity for language. And this capacity, I have argued, is not a function of the brain, but of the relationship between brains, signs, and the world. The capacity to interpret the sign symbolically is indeed the condition of human cognition. It remains only to minimize this assumption by providing a plausible context for the origin of the symbolic function.

There are, however, some specific constraints on a hypothesis of origin, beyond its plausibility and explanatory power for the interpretation of particular cultural and literary texts. In particular, the hypothesis must explain how the first “originary” sign was not a mere index but a symbol. It must explain, in short, the “paradox” of representing worldly objects via fictional or narrative signs.<sup>36</sup> The first symbol had to be understood already as a *meta-*

---

is it something we learn? There is no question that, *ontogenetically speaking*, we are innately predisposed to acquire language. But the difficulty of an evolutionary answer to the question of the origin of language is precisely that the genetic predisposition of modern *Homo sapiens* is itself a consequence of adaptation *to* symbolic reference (e.g., if modern *Homo sapiens* is innately predisposed to acquire language, the same cannot be said of hominid ancestors like *Homo erectus* and, a fortiori, *Australopithecus*). It follows that we cannot simply appeal to genetic factors in explaining the phylogenetic origin of an institution like language. The whole point of Deacon’s coevolutionary argument is that the origin of symbolic reference can only be explained as a response to socioecological pressures—which I think are also specifically anthropological pressures—and, furthermore, that these are ultimately only accessible in the form of a *deductive hypothesis* based on observations of both historical and contemporary human society. As Deacon (1997: 336) puts it: “I believe symbolic reference itself is the only conceivable selection pressure for such an extensive and otherwise counterproductive shift in learning emphasis [i.e., from index to symbol]. Symbol use itself must have been the prime mover for the prefrontalization of the brain in hominid evolution. Language has given rise to a brain which is strongly biased to employ the one mode of associative learning that is most critical to it.”

36. Searle (1979: 58) poses the problem of fiction as a paradox between sense and reference:

*phor* or *narrative projection* of its object, hence as qualitatively distinct from any previously given indexical sign. The referent of the first sign would be an instance of what Durkheim (1965 [1915]: esp. 225–45, 255–62) called the sacred. Henceforth the central object is understood only via the collective imposition of a metaphor designating its sacrality. This essentially negative or interdictive relationship makes the symbolic relationship between sign and object impossible to grasp indexically.

Is fiction or narrative really a consequence of ethical or religious interdiction? Is our remarkable facility with symbols a consequence of the deferral of sensorimotor or indexical response? Is the origin of symbolic reference essentially a process of negation, which is to say, of nascent irony, when we represent the opposite of what can be indexically inferred? From the perspective of iconic and indexical reference, the major obstacle to learning symbols is simply their counterintuitiveness. Why indeed try to indicate one thing by demonstrating its opposite? What explains the emergence of this type of reference?

Surveying the empirical evidence, Deacon (1997: chap. 11) argues that the dramatic expansion of the prefrontal cortex evident in the hominid fossil record from (roughly) late *Australopithecus* to archaic *Homo sapiens* is a result of selection for symbolic reference. The key to the prefrontal cortex, he argues, is that it functions as a mediator between direct sensory and motor functions of the brain. In particular, the prefrontal cortex is active in inhibiting “the tendency to act on simple correlative stimulus relationships,” thereby increasing the ability to sample “alternative higher order sequential or hierarchic associations” (ibid.: 265). The expansion of this particular area of the brain in human evolution suggests selection for the peculiar memory and learning requirements of symbolic thought.

Experiments with chimpanzees and other nonhuman primates have illustrated the general significance of the prefrontal cortex for overriding or inhibiting salient stimulus/stimulus and stimulus/response relationships in order to generate higher-order indexical associations. In one experiment cited by Deacon (1997: 261), food was placed in multiple wells while a monkey watched. The monkey was then allowed to retrieve the food. In the normal course of things, the monkey will not sample the same well twice. Once the food is removed, it moves on to the next well. But prefrontally damaged monkeys fail to sample efficiently; they perseverate, returning to the same well repeatedly, even after the food has been retrieved.

Deacon (1997: 264) suggests such experiments indicate the role of the pre-

---

How is it that fictional sentences can mean without referring? See van Oort 1998b for a critique of Searle’s logical model of fiction.

frontal cortex in generating higher-order indexical associations. In the case of the perseverating monkey, prefrontal damage prevents the animal from overriding the first rewarding stimulus in order to sample another well: it is as if the monkey cannot divorce the original stimulus (the well) from the motor response (the act of sampling the well). In the normal case, the original stimulus is of course remembered, but the original context is now precisely understood as indicating the *absence* of food. The ability to negate the original indexical relationship is, Deacon suggests, a crucial function of the prefrontal cortex.

These experiments give us some hints about the problem confronting a species that makes the leap from indexical to symbolic representation. The ability to work with highly abstract associations independently of immediately salient stimuli is the distinctive mark of symbolic thought. (Think, for example, of the ability to read a novel in a busy airport, or to rehearse in one's mind the main points of a lecture while driving a car, or to picture the perfect swan dive at the top of the five-meter diving platform.) The ease with which we do these things makes it difficult for us to grasp the challenge it presents to other species. But this is precisely the challenge that confronted our remote ancestors who originally made this leap from index to symbol. What peculiar circumstances could have led to such an event?

One particular experiment cited by Deacon (1997: 413–14) vividly illustrates the paradoxical conditions necessary for symbolic reference to get off the ground. In the first stage of the experiment, chimpanzees were required to choose between two unequal piles of candy. Not surprisingly, they always chose the larger pile. The experiment was then complicated when the selected pile was given not to the chimpanzee who chose it, but to another chimpanzee. Despite repeated trials, the chimps seemed unable to learn that the smaller pile should be selected in order to get the larger pile. In each trial, the chimpanzee watched in obvious frustration and disappointment as the larger pile it had selected was given to another. Deacon speculates that what makes the problem so difficult for the chimpanzee is the salience of the immediate stimulus. The chimp is unable to override this stimulus and conceptualize its opposite: choosing the smaller pile in order to get the larger. As Deacon (*ibid.*: 414) puts it, this solution is counter-intuitive for the chimp, because it “is overshadowed by the very powerful influence of its mutually exclusive and otherwise obvious alternative.” The difficulty lies in the paradox of selecting the lesser object in order to obtain the greater one.

Evidently, this type of problem is not habitually encountered by chimpanzees in their daily lives. Yet if Deacon is right, this kind of pragmatic paradox is a regular feature of human interaction. Deacon argues that it is

the peculiar learning bias granted by an expanded prefrontal cortex that makes such tasks both easy and natural for humans. But, he adds, the expanded prefrontal cortex is itself a consequence—not a cause—of the origin of symbolic reference. What circumstances could have led to the latter?

We see in these experiments an indirect example of the kind of “pragmatic paradox” that Eric Gans (1997: 13–36) proposes as the basis of symbolic representation.<sup>37</sup> The term *pragmatic paradox* is actually due to Paul Watzlawick, Janet Beavin, and Don Jackson (1967: 194–95), who have explored the pragmatic or behavioral consequences of paradox in communication, especially in pathological communicative situations that lead to schizophrenia. They are themselves elaborating on Gregory Bateson’s (1972: 201–27, 271–78) notion of the “double bind,” namely, the failure of a “victim” or schizophrenic patient to respond normally to a paradoxical statement from a caregiver or “mother,” owing to the victim’s inability to distinguish between different logical types or levels of abstraction in the meaning of the caregiver’s statement. Fearing rejection, the child is unable to draw the mother’s attention to the paradoxical nature of her statement and instead feels compelled to respond as if there were no paradox. This, Bateson argues, often leads to highly aberrant metaphoric interpretations and eventually, in extreme cases, to schizophrenia. For our present discussion, what is significant about Bateson’s double bind is that it locates the source of schizophrenia in a feature that, far from being abnormal, is actually very common to human communicative situations. Watzlawick, Beavin, and Jackson (1967: 199–200) multiply examples of pragmatic paradox, including commonplace injunctions like “Enjoy yourself!” or “Be spontaneous!”

Like Bateson, and like Watzlawick, Beavin, and Jackson, Gans sees pragmatic paradox as endemic to human communication, a basic feature of language. His interest in it, however, is not primarily psychiatric or psychological. Like Deacon, he argues that such paradoxes reveal a great deal about the evolutionary anomaly of language. In elaborating a hypothesis for language origin, we need to take into account both the anomalous nature of language compared to other animal communication systems and the paradoxical structure of symbolic reference itself.

Gans relates the two by hypothesizing that symbolic reference emerges as

37. For an introduction to Gans’s idea of generative anthropology, see the *Anthropoetics* Web site at [www.anthropoetics.ucla.edu](http://www.anthropoetics.ucla.edu). See also Wolfgang Iser (2000) on the different emphasis of his own “literary anthropology.” The conversation between Iser and Gans is continued in my interview with Iser (see van Oort 1998a) and in Gans 2000. For more literary applications of Gans’s work, see Schneider 1999 and Kohan 2001. Both provide excellent brief introductions to the development of mimetic theory in Girard and Gans.

a response to the pragmatic paradoxes engendered by presymbolic mimesis. Refining René Girard's (1965, 1977, 1987) theory of mimesis, Gans argues that imitation, which is the condition of all socially learned behavior, inevitably leads to conflict over scarce objects (such as food).<sup>38</sup> When B imitates A in order to appropriate object X, the situation produces conflict because A is no longer simply a model for B but an obstacle. As ethologists have noted,<sup>39</sup> mimetic conflict, which is potentially debilitating to the social order, is stabilized by the evolution of a pecking order; B defers to A's higher rank, leaving the object for the latter's possession. But, Gans points out, greater mimetic propensity makes this solution increasingly unstable. When B proceeds to imitate A by refusing to leave the scene of mimetic attraction, this creates a dangerous and inherently conflictive situation, which Gans (1997: 20) describes as a "pragmatic paradox" because what serves to increase the object's attraction also increases the danger of conflict. At a certain point, the mimetic symmetry of the situation undermines the asymmetry inherent to the pecking-order solution in which B defers to A or B overthrows A and assumes the alpha position. Instead, both A and B serve as mediators for one another: B imitates A and A imitates B. In this situation of mutual mimetic reinforcement of the other's desire for the object, neither A nor B can appropriate the object independently of the mediating and potentially dangerous presence of the other. But, equally, the high level of mimeticism prevents either individual from simply abandoning the scene.

This intolerable situation of pragmatic paradox is resolved when both A and B accept the other's abortive gesture of appropriation as no longer a straightforward index of the object's availability, that is, as a sign to be imitated in an unselfconscious strategy of object appropriation, but rather as a higher-order sign that negates the first-order indexical relation between sign and object. This higher-order sign refers only by negating the previous positive, indexical reference to the object. It is thus not truly an index at all. Its mode of reference is rather *paradoxical* because it succeeds only insofar as it undermines the salience of the self-evident connection between gesture and object. The originary sign thus signifies by *negating* each individual's prior indexical attachment to the object. The consequence is a dynamic

38. For a different (but nonetheless highly provocative) approach to the origin of language, see Donald 1991, 1999. Donald (1999: 148) is more concerned with the cognitive basis of presymbolic mimesis than with its conflictive social consequences, but he does note that, as an evolutionary strategy, the emergence of mimesis "might also have introduced some destabilizing elements, especially by amplifying both the opportunities for competition, and the potential social rewards of competitive success."

39. Jane Goodall's *In the Shadow of Man* (1988 [1971]) is the classic ethological study of wild chimpanzees, but see also Franz de Waal (1998 [1982]). Both accounts are vivid and highly engrossing depictions of chimpanzee social life.

oscillation between sign and referent, in which the sign oscillates between the dual modes of index and symbol. As an index, the sign indicates the presence of the object, but this concrete reference to the object is immediately punctured by the realization that the sign is equally a response to the other's symmetrically produced sign. The individual abandons the original indexical conception of the sign and returns to a nascent symbolic understanding, which, as Deacon persuasively shows, functions not by referring directly to the object, but by mediating this reference to the object through other signs. The pertinence of Gans's hypothesis is that it allows us to understand that the origin of this new sign/sign relation is none other than the mimetic relation between subject and rival, which is to say, ultimately, the relation between the individual and the entire sign-using community. Durkheim (1965 [1915]) had already understood that the origin of religion lies in the projection of the social relation onto totemic objects, but he could give no further account of projection as a function of the symbolic sign.

The originary sign thus solves the problem of maintaining a highly salient stimulus in view while simultaneously refusing to act on this same stimulus. This is precisely the problem (i.e., being presented with a pile of candy that could only be appropriated by negating the normal response pattern) that Boysen's chimpanzees in the candy experiment could not solve. In the experiment, the chimp fails not simply to defer immediate gratification but to *project* the salience of the primary indexical relation onto a parallel symbolic reality, in which reference is constituted, not by physical or worldly co-occurrence but by the relationships between symbols. As Deacon implies and Gans explicitly hypothesizes, the originary model for the projection of an indexical sign to a symbolic sign is negation. In negating the priority of the indexical relation, the sign refers to its object only by paradoxically representing its opposite. The symbolic sign is thus fundamentally ironic; it points to its object only to reject it. In effect, the sign says "here is the object, but you cannot have it."

The paradoxical structure of symbolic reference is similar, but not identical, to that of play, as Bateson (1972: 177-93) has analyzed it. In Bateson's account, what separates play fighting from real fighting is the frame of reference: the participants in the former are distinctively self-conscious about their movements. Bateson (*ibid.*: 179) is keen to see in this transition to "metacommunication" (i.e., from signals to an awareness that "signals are signals") a major evolutionary step that implies not only "language . . . but also all the complexities of empathy, identification, projection and so on." Implied in Bateson's analysis of play, and of paradoxical communicative structures in general, is the transition from indexical to symbolic reference. In light of Deacon's neurological and primatological research, how-



ever, it seems doubtful that Bateson's (ibid.) casual observations of monkeys playing in the zoo will bear the weight of his claim that they already demonstrate the paradoxical structure of symbolic reference. After all, even genuine fights for dominance are highly "ritualized" or stereotyped affairs that depend largely on the proficiency of the signaler to communicate his strength without actually having to come to blows. Only with a hypothesis for the emergence of symbolic reference can we begin to analyze the paradoxical structure inherent to such human practices as play, ritual, irony, and art.

Let us take the case of irony.<sup>40</sup> The standard definition of *irony* is saying the opposite of what one really means: a "figure of speech in which the intended meaning is the opposite of that expressed by the words used" (*OED*). What the standard definition captures is the fact that irony functions as a negation of a preexisting meaning. But this negation is not accomplished simply by adding *not* to a previous statement, as with logical negation. Rather, ironic negation functions by reframing or recontextualizing the existing sense of the phrase or sentence. The meaning is not simply canceled but situated in a new (ironic) context, so that a tension is generated between what is said and what is really the case.

But this tension between meaning and reality, far from being a mere figure of speech, is in fact the *sine qua non* of symbolic reference. As we have seen, symbolic reference emerges as the negation of its independent indexical constituents, which nonetheless provide the basis for interpreting the sign symbolically. Rather than respond on the basis of a previously given perceptual and motor pattern, the subject of pragmatic paradox responds by accepting the indexical sign *as a sign of itself*. It is this paradoxical attention to the *form* of the sign-in-itself that undermines the naturalness inherent to the indexical response, which interprets the sign/object relation spatio-temporally as a relation of part to whole or part to part. This relation is something that is most often directly perceived, as in the classic example of smoke (the sign) indicating fire (the object). Even in the case of "arbitrary" alarm calls, such as those produced by vervet monkeys, the interpretation of the relationship between sign (alarm call) and object (predator) is still dependent upon the assimilation of the arbitrary structure of the sign/object relation to the indexical associations of direct perceptual experience.

In the situation of pragmatic paradox, on the other hand, the subject's attention is drawn away from the object itself toward the object's media-

40. My analysis is indebted to Gans's (1997: 64–74) illuminating discussion of irony in *Signs of Paradox*.

tion by the indexical sign. The index becomes a meta-index, that is, a sign that thematizes its relation to the object, a relation now understood to be a consequence not simply of predictable co-occurrence, but of the meaning or “being” granted to the object by the co-participants in the symbol-using community. What is specifically ironic in the symbolic relationship is the sign’s irreconcilable difference from the indexical mode of interpretation. The sign is linked to its object not naturally, but conventionally: in the former, sign and object coexist in a “horizontal” spatiotemporal relationship; in the latter, the sign is pursued as a response not to the object, but to the other’s sign. Thus, once the sign is divorced irrevocably from its indexical object, its “meaning” can itself become the object of a new ironic meaning. The arbitrariness or “verticality” (Gans 1997: 14) of the signifier is a corollary of this fact. When the child imitates its parent who utters the word *cat* while pointing to a real cat, the child learns the meaning of the sign ostensively. But once the meaning of the word has been learned, it can equally be applied to other cats. In learning the word *cat*, the child does not seek to pursue the ostensive referent; it learns to *name* the cat by imitating the word produced by the parent. In the ontogeny of language acquisition, the conflict of appropriative mimesis (Girard 1977, 1987) is wholly eliminated. What the child imitates is not the nonlinguistic gesture toward the object, which leads to conflict, but the signification of the object. Unlike the object it designates, the word may be easily reproduced and thus universally shared. In the symbolic configuration that defines human communication, the imitation of the sign can be successfully pursued precisely because the goal of imitation is no longer the appropriation of the object but rather a self-conscious thematization of it.

It would, however, be a gross oversimplification to see this as a case of ontogeny recapitulating phylogeny. The child’s language acquisition is biased from the start thanks to hundreds of thousands—perhaps even millions—of years of brain-language coevolution. It has therefore become all but “natural” for the child to imitate gestures of meaning rather than of appropriation, symbols rather than indices. But in phylogenetic terms, we cannot assume the “naturalness” or “givenness” of this easy transition. This is the major conclusion reached by Deacon (1997: 36), on independent empirical grounds, in his study of primate brain evolution. And in presenting a minimal originary hypothesis for language origin, Gans (1997) accepts the necessity of this conclusion. The transition from index to symbol, from “natural” to “conventional” signification<sup>41</sup> cannot be discerned

41. For a painstaking philosophical discussion of the dualism implied by the distinction between natural and conventional meaning, see Rollin 1976. I share Rollin’s (1976: 86) desire to reduce this (metaphysical) dualism to a (pragmatic) distinction, but I am skeptical of his

empirically, through stages of increasingly complex indexical and iconic reference strategies. But nor should it be left as a metaphysical conundrum that can only be interpreted in terms of an irreconcilable ontological dualism (e.g., idealism vs. materialism, conventional vs. natural meaning, nature vs. culture, etc.). Rather, the difference should be minimized in terms of a hypothesis of origin. According to Gans's (1997) hypothesis, the pragmatic paradox precipitated by the presymbolic mimetic relationship between subject and model is resolved when each individual learns to pursue the sign not as an index but as a conventional sign. It goes without saying that this conventional relationship can always be undermined or deconstructed by pointing out the ironic dependence of the sign on its indexical constituents.<sup>42</sup>

Irony defines the mode of attention paradigmatic of symbolic interpretation, because it demands that we continually seek to counter the reduction of a symbol to an index. What tempts us toward irony is the feeling that our interlocutor knows how we will respond based on what we can both perceive to be the case. If I say, "What a gorgeous day for a walk!" in the face of torrential rain, I am speaking ironically because the rain is an index that both I and my interlocutor associate with the unpleasant experiences of getting wet, being cold, and so on. My words are not an index of what is really the case, but of what I would like to be the case. In short, I project onto the real situation of rain an imaginary situation in which it is sunny. In referring to the imaginary situation, I compensate my resentment toward the real situation of a torrential downpour. We are attuned to the aesthetics of irony because it captures the fundamental incompatibility between the symbolic world we desire and the indexical world in which we live. Whether we are commenting ironically on the weather or appreciating the ironies of Greek tragedy, the same fundamental structure operates. In each case we are displacing an unpleasant indexical experience onto a symbolic world in which that experience becomes an occasion for pleasurable irony.<sup>43</sup>

---

appeal to psychology as an adequate answer to the origin of the distinction. In particular, I do not share Rollin's (*ibid.*: 97) rather uncritical acceptance of "the principle that ontogeny recapitulates phylogeny" for precisely the same reasons that I do not accept Lakoff and Johnson's reduction of meaning to perceptual association.

42. This analysis suggests a solution to the logical paradoxes generated by metaphysical discussions of ostension or pointing, which seek to reduce the latter to either an exclusively natural or conventional phenomenon (see Rollin 1976: 99–100). Meir Sternberg (1999: 372ff.) points out that the traditional metaphysical dichotomy between arbitrary and natural signs can be more usefully understood in terms of the Peircean trichotomy of icon, index, and symbol, with the index reconceived. Sternberg's point, I take it, is rather similar to Deacon's. Indexical reference provides the crucial missing link in rethinking the traditional dualisms between arbitrary and natural signification systems.

43. Irony is also the source of many a joke, including the following one that highlights the dif-

From our external, historically privileged perspective, it is perhaps too easy to understand the original projection from index to symbol as an act of ironic negation of indexical relations. From the vantage point of modern symbol users, we immediately grasp the utility of substituting a symbolic object for the real one and account for this utility by pointing to the greater referential economy permitted by language in the long run. But at the moment of origin, no such perspective is possible. To the original participants, the inaccessibility of the central object cannot be grasped in terms of a “theory of projection.” Symbolic projection is then rather understood as a revelatory event, upon which the very survival of the community depends. According to Gans (1993: 117), this provides the originary model of the sacred; it captures the fact that sacrality is not in the first place a naive fantasy produced by overimaginative minds—which must in fact be the conclusion of Turner’s cognitive model, were it applied to the sacred. Rather, the sacred stands at the origin of projection. In projecting the deferring function of the sign onto the central object, the community interprets the object as the cause of this deferral. The object is no longer an index of an appetitive need but a symbol or figure representing the collective interdiction of all such needs. And the most powerful collective symbol is the divine being or god who alone is capable of withstanding the multiple and conflicting desires of the surrounding community.

This model of “originary projection” also explains the fundamental distinction between true symbolic narrative and the mere indexical association

---

ference between indexical and symbolic thought in a way that nicely complements Boysen’s candy experiment with chimpanzees. Two children, Max and Ruby, are offered two slices of cake. One piece, however, is significantly larger than the other. “You choose!” says Ruby, knowing that, by committing Max to the first choice, she has also committed him to taking the smaller piece as politeness demands. To her dismay, Max instead helps himself to the larger piece. Indignant, Ruby attempts to draw his attention to the infraction:

Ruby: “Why did you take the larger piece?”

Max: “Which piece would you have taken?”

Ruby: “Why, the smaller of course.”

Max: “But that is what you have! Why are you now upset?”

The joke plays on the irony of substituting signs for things. Both Max and Ruby are of course primarily concerned with getting the biggest piece of cake. But since both are also competent language users, each knows that the other’s appetite is mediated by symbolic rules that dictate how desirable objects (like cake) should be appropriated. In this case, politeness requires that the smaller piece be taken when offered the choice. Knowing this, Ruby offers the choice to Max. But Max surprises her, and us, by exploiting the irony latent in all such rules. If Ruby really meant what she said, she should be content with the smaller piece. The joke plays on the fact that it is not sincerity that is important but merely civility. Sincerity, as the joke demonstrates, is the alibi used by the romantic who refuses to play by the rules of normative ethical conduct. (My thanks to Marina Ludwigs for bringing my attention to this joke, whose original version can be found in Astrid Lindgren’s *Karlsson on the Roof*.)

of contiguous events. Narrative implies a symbolically constructed agent or self (which is, in the first place, a god or divine agent, not a human—the latter rather being modeled on the former). The ability simply to perceive and remember sequences of contiguous events, however, entails no additional effort to understand this sequence as produced by an agent or self different from the perceiving subject.<sup>44</sup>

Ritual and myth commemorate the origin of symbolic reference by attributing the deferring function of representation not just to the sign, but to an entire numinous ontology of the sacred. From our perspective, this is a category error, an expansion or projection of the symbolic world to the natural world. But this original “alienated” mode of symbolic experience is not something that can be eliminated once and for all. It can only be *minimized* by recognizing its proper place in anthropological and historical explanation.

If today we understand the conflict-deferring function of ritual and myth as the minimal anthropological origin of the symbolic sign, then to the original participants no such self-understanding is possible, because no such level of historical self-reflection is available at the very first moment of history. Our ability to formulate a hypothesis of cultural-symbolic origin is a condition of our historical distance from this origin. In taking this entire history for granted, the cognitive model of representation reduces what is only accessible through concrete anthropological and historical analysis to an ontology of the brain and its neurological connections. This is of course consistent with the scientific and metaphysical ideal since René Descartes. But, as Deacon’s work makes clear, it is not consistent with a more subtle account of the coevolution of language and the brain. And the simplest demonstration of this truth, I have argued, is the fact that language—symbolic reference—is a specifically human mode of representation, qualitatively distinct from the indexical systems that preceded it.

What are the practical consequences of my analysis for literary study? I am not saying that we can ignore cognitivism. On the contrary, I think that this

44. Implied in this hierarchical shift from indexical reference by sequential association to true narrative representation is the presence of what psychologists call “theory of mind” (see Dunbar 1998: 102). By this label psychologists mean not a philosophy of mind, but simply the ability to judge other people’s mental states and hence to distinguish between one’s own beliefs and those of others. Intentionality or “theory of mind” may be multiplied to quite high levels, e.g., I believe that X believes that Y believes that X is sleeping with Y’s partner. In the case of fictional narratives, this layering is further complicated by the fact that all the intentional relationships between characters are subordinated to our sense of an overarching authorial presence, a sort of meta-intentionality that is not simply identical with the intentionality of the real author.

kind of cross-disciplinary research is exciting and necessary. At the same time, this research needs to be more fully aware of some of the traditional philosophical problems involved in any account of the origin of representation. More precisely, I do not think it is possible in our “postmetaphysical” era to assume that symbolic representation can be traced so easily to internal ideas or perceptions, in a way fundamentally no different than John Locke’s. No doubt, with modern developments in evolutionary biology and developmental psychology, cognitivism is far better equipped to grasp the essential uniformity of human experience. Unlike Locke (1975 [1690]: 104), who famously used the metaphor of the mind as a sheet of “white paper,” biology does not need to rely on pure metaphor for an explanation of the mind as a product of evolution (if it did, then biology would be a poetics rather than a science). But when it comes to the problem of the origin of symbolic representation, the cognitivist model is no better off than Locke’s notion of the *tabula rasa*. And this is ultimately because the mind, whether as white paper or as a biologically evolved adaptive system, cannot construct symbols without the anthropological context of the originary *scene* of representation.

Instead of taking the presence of such a scene for granted, I believe we need to confront the problem of representation head on. The originary hypothesis explains the origin of symbolic representation not as a *consequence* of biological evolution or the struggle for reproductive fitness, but indeed as a new system that coevolves with biology. The general semiotic approach I advocate requires that we trace and explain the shift from an indexical to a symbolic representational system. Both Deacon and Gans, upon whom I have relied for much of this argument, understand the problem of human origin to be a problem about the origin of symbolic reference. But I do not want to give the impression that they have said the last word on representation. Far from it. The single most important point I wanted to make is this: If we want to explain the origin of what we take to be universal human categories, for example, metaphor and narrative, then we are going to have to bite the bullet and explain what the anthropological function of these categories is. But in order to do that, we need to take seriously the fact that symbolic representation is irreducible to the causal mechanisms of biological functions. Instead, positively speaking, we need to recognize that the only explanations available to us lie in the functioning of the historical institutions of human culture itself. This does not mean that we have to renounce the search for a unified explanation. On the contrary, the whole point of the originary hypothesis is to minimize our explanation of culture, and all of its many forms, by presenting a minimal anthropological context from which they can be more generally understood. To put it as pithily as

possible: the answer to biological reductionism is not cultural idealism, but anthropological mimimalism.

## References

- Bateson, Gregory  
1972 *Steps to an Ecology of Mind* (San Francisco: Chandler).
- Bickerton, Derek  
1990 *Language and Species* (Chicago: University of Chicago Press).  
1995 *Language and Human Behavior* (Seattle: University of Washington Press).
- Burling, Robbins  
1993 "Primate Calls, Human Language, and Nonverbal Communication," *Current Anthropology* 34 (1): 25-53.
- Calvin, William  
1997 "Talking Heads" (review of Terrence Deacon, *The Symbolic Species*), *New York Times on the Web*, 10 August. Accessed on-line (17 October 2000) at [www.nytimes.com/books/97/08/10/reviews/970810.10calvint.html](http://www.nytimes.com/books/97/08/10/reviews/970810.10calvint.html).
- Calvin, William, and Derek Bickerton  
2000 *Lingua Ex Machina: Reconciling Darwin and Chomsky with the Human Brain* (Cambridge: MIT Press).
- Cassirer, Ernst  
1944 *An Essay on Man: An Introduction to a Philosophy of Human Culture* (New Haven, CT: Yale University Press).
- Chomsky, Noam  
1972 *Language and Mind* (New York: Harcourt Brace Jovanovich).
- Crane, Mary Thomas, and Alan Richardson  
1999 "Literary Studies and Cognitive Science: Toward a New Interdisciplinarity," *Mosaic* 32 (2): 123-40.
- Dawkins, Richard  
1989 [1976] *The Selfish Gene* (Oxford: Oxford University Press).
- Deacon, Terrence  
1997 *The Symbolic Species: The Co-Evolution of Language and the Brain* (New York: Norton).
- Dennett, Daniel C.  
1995 *Darwin's Dangerous Idea: Evolution and the Meanings of Life* (New York: Simon & Schuster).
- Derrida, Jacques  
1974 *Of Grammatology*, translated by Gayatri Spivak (Baltimore, MD: Johns Hopkins University Press).
- Donald, Merlin  
1991 *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition* (Cambridge: Harvard University Press).  
1998 "Mimesis and the Executive Suite: Missing Links in the Origin of Language," in *Approaches to the Evolution of Languages: Social and Cognitive Bases*, edited by James Hurford, Michael Studdert-Kennedy, and Chris Knight, 44-67 (Cambridge: Cambridge University Press).  
1999 "Preconditions for the Evolution of Protolanguages," in *The Descent of the Mind: Psychological Perspectives on Hominid Evolution*, edited by Michael C. Corballis and Stephen E. G. Lea, 138-54 (Oxford: Oxford University Press).
- Dunbar, Robin  
1998 "Theory of Mind and the Evolution of Language," in *Approaches to the Evolution of Languages: Social and Cognitive Bases*, edited by James Hurford, Michael Studdert-Kennedy, and Chris Knight, 92-110 (Cambridge: Cambridge University Press).

- Dunbar, Robin, Chris Knight, and Camilla Power, eds.  
1999 *The Evolution of Culture: An Interdisciplinary View* (New Brunswick, NJ: Rutgers University Press).
- Durkheim, Emile  
1965 [1915] *The Elementary Forms of the Religious Life*, translated by Joseph Ward Swain (New York: Free Press).
- Englefield, F. R. H.  
1977 *Language: Its Origin and Its Relation to Thought*, edited by G. A. Wells and D. R. Oppenheimer (London: Pemberton).
- Fischer, Michael  
1996 "Literary Change and Cognitive Science" (Review of Spolsky, *Gaps in Nature: Literary Interpretation and the Modular Mind*), *Poetics Today* 17: 362–65.
- Fish, Stanley  
1980 *Is There a Text in This Class?* (Cambridge: Harvard University Press).
- Fludernik, Monika, Donald C. Freeman, and Margaret H. Freeman  
1999 "Metaphor and Beyond: An Introduction," *Poetics Today* 20: 383–96.
- Fox, Robin  
1980 *The Red Lamp of Incest* (London: Hutchinson).  
1983 *Kinship and Marriage: An Anthropological Perspective* (Cambridge: Cambridge University Press).
- Gans, Eric  
1993 *Originary Thinking: Elements of Generative Anthropology* (Stanford, CA: Stanford University Press).  
1997 *Signs of Paradox: Irony, Resentment, and Other Mimetic Structures* (Stanford, CA: Stanford University Press).  
1999 "The Little Bang: The Early Origin of Language," *Anthropoetics—The Electronic Journal of Generative Anthropology* 5 (1). Accessed 15 April 2002 at [www.anthropoetics.ucla.edu/apo501/gans.htm](http://www.anthropoetics.ucla.edu/apo501/gans.htm).  
2000 "Staging As an Anthropological Category," *New Literary History* 31: 45–56.
- Geertz, Clifford  
1973 *The Interpretation of Cultures* (New York: Basic Books).
- Genette, Gérard  
1980 *Narrative Discourse: An Essay in Method*, translated by Jane E. Lewin (Ithaca, NY: Cornell University Press).
- Girard, René  
1965 *Deceit, Desire, and the Novel*, translated by Yvonne Freccero (Baltimore, MD: Johns Hopkins University Press).  
1977 *Violence and the Sacred*, translated by Patrick Gregory (Baltimore, MD: Johns Hopkins University Press).  
1987 *Things Hidden since the Foundation of the World*, translated by Stephen Bann and Michael Metteer (Stanford, CA: Stanford University Press).
- Goldschmidt, Richard  
1952 "Evolution, As Viewed by One Geneticist," *American Scientist* 40: 84–135.
- Good, Graham  
2001 *Humanism Betrayed: Theory, Ideology, and Culture in the Contemporary University* (Montreal and Kingston: McGill-Queen's University Press).
- Goodall, Jane  
1988 [1971] *In the Shadow of Man*, rev. ed. (Boston: Houghton Mifflin).
- Gross, Sabine  
1997 "Cognitive Readings: Or, the Disappearance of Literature in the Mind," *Poetics Today* 18: 271–97.
- Harris, Marvin  
2001 [1968] *The Rise of Anthropological Theory: A History of Theories of Culture*, updated ed. (Walnut Creek, CA: Altamira).



- Hauser, Matt D.  
1996 *The Evolution of Communication* (Cambridge: MIT Press).
- Herman, David  
1999 "Parables of Narrative Imagining," *Diacritics* 29: 20–36.
- Hernadi, Paul  
2001 "Literature and Evolution" *SubStance* 94/95: 55–71.
- Hobbes, Thomas  
1947 [1651] *Leviathan*, edited by Michael Oakeshott (Oxford: Basil Blackwell).
- Hockett, Charles F., and Robert Ascher  
1964 "The Human Revolution," *Current Anthropology* 5 (3): 135–68.
- Holloway, Ralph  
1969 "Culture: A Human Domain," *Current Anthropology* 10 (4): 395–412.  
1998 "Language's Source: A Particularly Human Confluence of Hard Wiring and Soft" (review of Terrence Deacon, *The Symbolic Species*), *American Scientist* 86: 184–86.
- Hudson, Richard  
2001 Review of Terrence Deacon, *The Symbolic Species*, *Journal of Pragmatics* 33: 129–35.
- Hurford, James  
1998 Review of Terrence Deacon, *The Symbolic Species*, *Times Literary Supplement*, October 23, 32.
- Hurford, James, Michael Studdert-Kennedy, and Chris Knight, eds.  
1998 *Approaches to the Evolution of Languages: Social and Cognitive Bases* (Cambridge: Cambridge University Press).
- Iser, Wolfgang  
2000 "What Is Literary Anthropology? The Difference between Explanatory and Exploratory Fictions," in *Revenge of the Aesthetic: The Place of Literature in Theory Today*, edited by Michael Clark, 157–79 (Berkeley: University of California Press).
- Jackson, Tony  
2000 "Questioning Interdisciplinarity: Cognitive Science, Evolutionary Psychology, and Literary Criticism," *Poetics Today* 21: 319–47.
- Johnson, Mark  
1987 *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason* (Chicago: University of Chicago Press).
- Knight, Chris  
1991 *Blood Relations: Menstruation and the Origins of Culture* (New Haven, CT: Yale University Press).
- Kohan, Kevin  
2001 "James and the Originary Scene," *Henry James Review* 22: 229–38.
- Lakoff, George  
1987 *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press).
- Lakoff, George, and Mark Johnson  
1980 *Metaphors We Live By* (Chicago: University of Chicago Press).  
1999 *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought* (New York: Basic Books).
- Lakoff, George, and Mark Turner  
1989 *More Than Cool Reason: A Field Guide to Poetic Metaphor* (Chicago: Chicago University Press).
- Locke, John  
1975 [1690] *An Essay concerning Human Understanding*, edited by Peter H. Nidditch (Oxford: Oxford University Press).
- Lodge, David  
1977 *The Modes of Modern Writing: Metaphor, Metonymy, and the Typology of Modern Literature* (Chicago: University of Chicago Press).

- Miers, Paul  
1992 "The Other Side of Representation: Critical Theory and the New Cognitivism," *Modern Language Notes* 107: 950-75.
- Miller, Geoffrey F.  
1999 "Sexual Selection for Cultural Displays," in Dunbar et al. 1999: 71-91.
- Müller, Max  
1873 *Introduction to the Science of Religion* (London: Longmans).
- Peirce, Charles Saunders  
1956 [1897, 1903] "Logic As Semiotic: The Theory of Signs," in *The Philosophical Writings of Peirce*, edited by J. Buchler, 98-119 (New York: Dover).
- Pinker, Steven  
1994 *The Language Instinct* (New York: HarperCollins).
- Richardson, Alan  
1999 "Cognitive Science and the Future of Literary Studies," *Philosophy and Literature* 23: 157-73.
- Rollin, Bernard E.  
1976 *Natural and Conventional Meaning: An Examination of the Distinction* (The Hague: Mouton).
- Rousseau, Jean-Jacques  
1994 [1755] *Discourse on the Origin of Inequality*, translated by Franklin Philip and edited by Patrick Coleman (Oxford: Oxford University Press).
- Ruse, Michael  
1998 "The Evolution of Symbols" (review of Terrence Deacon, *The Symbolic Species*), *Semiotic Review of Books* 9 (3). Accessed 19 February 2001 at [www.chass.utoronto.ca/epc/srb/srb/evolution.html](http://www.chass.utoronto.ca/epc/srb/srb/evolution.html).
- Sampson, Geoffrey  
1997 *Educating Eve: The "Language Instinct" Debate* (London: Cassell).
- Savage-Rumbaugh, Sue E.  
1986 *Ape Language: From Conditioned Response to Symbol* (New York: Columbia University Press).
- Schneider, Matthew  
1999 "Problematic Differences: Conflictive Mimesis in Lessing's *Laokoon*," *Poetics Today* 20: 273-89.
- Searle, John  
1979 *Expression and Meaning: Studies in the Theory of Speech Acts* (Cambridge: Cambridge University Press).  
1992 *The Rediscovery of Mind* (Cambridge: MIT Press).  
1995 *The Construction of Social Reality* (New York: Free Press).
- Sebeok, Thomas  
1994 *Signs: An Introduction to Semiotics* (Toronto: University of Toronto Press).
- Seyfarth, Robert M., Dorothy Cheney, and Peter Marler  
1980 "Monkey Responses to Three Different Alarm Calls: Evidence of Predator Classification and Semantic Communication," *Science* 210: 801-3.
- Sternberg, Meir  
1999 "The *Laokoon* Today: Interart Relations, Modern Projects and Projections," *Poetics Today* 20: 291-379.
- Tooby, John, and Leda Cosmides  
2001 "Does Beauty Build Adapted Minds? Towards an Evolutionary Theory of Aesthetics, Fiction and the Arts," *SubStance* 94/95: 6-27.
- Turner, Mark  
1991 *Reading Minds: The Study of English in the Age of Cognitive Science* (Princeton, NJ: Princeton University Press).  
1996 *The Literary Mind* (New York: Oxford University Press).  
1998 "Poetry for the Newborn Brain" (review of Terrence Deacon, *The Symbolic Species*), *Bostonia* 1: 72-73.

- Turner, Mark, and Gilles Fauconnier  
 1999 "A Mechanism of Creativity," *Poetics Today* 20: 397–418.
- van Oort, Richard  
 1998a "The Use of Fiction in Literary and Generative Anthropology: An Interview with Wolfgang Iser," *Anthropoetics — The Electronic Journal of Generative Anthropology* 3 (2). Accessed 15 April 2002 at [www.anthropoetics.ucla.edu/apo302/Iser\\_int.htm](http://www.anthropoetics.ucla.edu/apo302/Iser_int.htm).  
 1998b "Three Models of Fiction: The Logical, the Phenomenological, and the Anthropological (Searle, Ingarden, Gans)," *New Literary History* 29: 439–65.
- Waal, Frans B. M. de  
 1998 [1982] *Chimpanzee Politics: Power and Sex among Apes*, rev. ed. (Baltimore, MD: Johns Hopkins University Press).
- Watzlawick, Paul, Janet Helmick Beavin, and Don D. Jackson  
 1967 *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes* (New York: Norton).
- White, Leslie A.  
 1949 *The Science of Culture* (New York: Farrar, Straus and Cudahy).  
 1959 *The Evolution of Culture: The Development of Civilization to the Fall of Rome* (New York: McGraw-Hill).  
 1987 *Ethnological Essays*, edited by Beth Dillingham and Robert L. Carneiro (Albuquerque: University of New Mexico Press).
- Wilson, Edward O.  
 1978 *On Human Nature* (New Haven, CT: Yale University Press).  
 1999 *Consilience: The Unity of Knowledge* (New York: Random House).