Integrating Indexicals in Simian Semiotics: Symbolic Development and Culture

Seth Surgan

Clark University and Eastern Connecticut State University

and

Simone de Lima

Clark University and Universidade de Brasília

The ability to understand both the self and others as purposeful agents — with thoughts, beliefs, and desires — seems to be central to the emergence of cultural processes both phylo- and ontogenetically. This ability has been termed second-order intentionality or "theory of mind" and has been conceptualized as a species-specific "trait" which is genetically predetermined, naturally selected and the resident of a dedicated module (i.e., a functional subsystem with an evolutionary history more-or-less independent of other such subsystems) within the mind. Alternatively, we see it emerging out of a more general process — symbolization. The paper discusses the emergence of the symbolic function from previously existing forms of communication by analyzing the structures and functions of different kinds of signs used in human and non-human vocal communication. We reinterpret evidence from the study of non-human primate vocalizations and suggest that these vocalizations embody a semiotic type that, like all signs, is more highly developed than a signal, but is not catalogued within the basic Peircean triad of sign types (i.e., icon, index, symbol). This form, the double indexical, is intermediary between indexes and symbols. We speculate on what structural and functional reorganization is required to establish a developmental continuity from signals through the various types of signs (including the double indexical) to the well-known structure of the symbol — and possibly beyond.

Culture has freed evolution from determination by direct (physical) organism/environment interaction and has allowed for an unprecedented level of flexibility in human ecology without the need for corresponding anatomical change. Along with its conceptual cousins, "tool use" and "language," "culture" has been at the center of discussions of the evolutionary continuities and discontinuities that simultaneously relate and differentiate humans and other primates. However, the debate over "who has what" too often concentrates on (dis)similarities between human and non-human primates and tends toward the construction of static (dis)continuities and typologies. Despite the popularity of claims that primates and humans are either the same or not, those claims shed little light on the genetic (i.e., developmental) relation between the various phenomena observed in these species. We would like to pursue a different path. By taking a comparative–genetic perspective, we will try to elaborate the general developmental processes that give rise to a wide range of communicative phenomena observable across both phylogenesis and ontogenesis.

In this paper, we offer a set of ideas about the co-development of symbolic communication and the processes through which other minds are interpreted. We do not want to tell another natural selection "just-so" story about the emergence of symbols or the ability to understand other minds. Instead, we will draw from semiotic and basic developmental theories to investigate the potentials and constraints of communicative systems used by human and non-human primates and their roles in the co-evolution of semiotic and intentional processes.[1]

Debating Primate Cultures

The discovery of primate behavioral traditions has spurred theorists in their debate over the status of primate culture (Nishida, 1986; Tomasello, 1994). Specifically, forms of behavior such as tool use (e.g., termite "fishing" [McGrew, 1992]), gestural ritualization [Kummer, 1971], food preparation (e.g., potato and wheat washing [Itani, 1958; Kawai, Tsumori, and Motoyoshi, 1963; Kawamura, 1959]), and play (Huffman, 1984) are well documented in the primatological literature. Some researchers have claimed that such forms of behavior warrant the label of "culture" since they (1) differ across groups of the same species, (2) are not strictly conditioned by the environment, and (3) are not a function of genotypic change (Boesch, 1993; Kawamura, 1959; McGrew, 1992; Nishida, 1986).

Tomasello (1994, 1998, 1999) argues that not all ontogenetically acquired, group-specific behavioral traditions should be called "culture" because they do not exhibit cumulative cultural evolution or what he terms "the ratchet effect." For Tomasello, the ratchet effect is the product of cultural learning, a

---

[1]Understanding the mental states of others implies understanding others as intentional beings, as will be clarified below.

specifically human mode of social learning (Tomasello, Kruger, and Ratner, 1993). Cultural learning differs from other types of social learning in that the former takes advantage of the capacity to consider others as mental agents. Taking others as mental agents allows the interpretation of others' actions as being guided by thoughts, desires, and beliefs (i.e., intentional states — Dennett, 1971, 1987, 1996; Tomasello, 1999).[2] The reconstruction of others' mental states enables humans to learn from each other in a way that leads to a high fidelity of transmission (Tomasello, 1994, p. 313), or exact reproduction of both the form and function of a behavior. Such reproduction provides the basis for younger generations to make subsequent innovations, thus constituting the "ratcheting" of cultural products. Although we find Tomasello's emphasis on reproduction or "faithful transmission" of cultural products problematic (de Lima and Surgan, in press), we would like to take a second look at — and critically discuss — his claim that understanding others as intentional is a prerequisite for entrance into the cultural world.

Intentionality and Explaining Behavior

Mental states (beliefs, desires, etc.) can be understood as one type of intentional phenomena. Intentional phenomena, in general, are *about* some other thing (Dennett and Haugland, 1987). One possible way of explaining behavior is by assuming that the entity producing the behavior is an *intentional system* — i.e., one capable of having mental states of some kind (e.g., beliefs or desires, which are necessarily *about* something) — and then formulating explanations that utilize those constructs (e.g., He did X because he *wanted* Y). The notion of "levels of intentionality" allows us to approach the question of what kinds of beliefs those might be (Dennett, 1987). Whereas an explanation of behavior based on zero-order intentionality does not attribute beliefs, an explanation utilizing first-order intentionality makes the claim that action was motivated by belief. Second-order intentionality, in contrast, entails beliefs about beliefs (i.e., beliefs about one's own beliefs as well as the beliefs of another).

Understanding others as purposeful agents — whose acts can be interpreted and tentatively predicted by using constructs such as desires, thoughts, and beliefs — qualitatively transforms the perception and interpretation of the world. For example, explanations of behavior based on higher-order intentionality transcend analysis of means–ends couplings (e.g., "Remove top of jar for cookie"). Reconstructing the mentally mediated aspects of action (e.g., "She opened the cookie jar to get a cookie") plays an important role in understanding the relation between behavior and its results.

---

[2]Following Premack and Woodruff (1978), this capacity has been referred to as "theory of mind" by researchers in developmental psychology and primatology (Baron–Cohen, 1995; Bartsch and Wellman, 1995; Carruthers and Smith, 1996; Meltzoff and Gopnick, 1993; Povinelli, 1996).

A grasp of others' intentional states, according to Tomasello, helps the individual enter the world of cultural meanings and conventions. However, the minds of others are not immediately perceived. Developing knowledge about others as intentional, mental beings, therefore, necessarily requires a representational process. Given that many different forms of representation are conceivable, it seems important at this point to look into the nature of the representational processes that may serve this purpose.

Representing Minds

What form of representation would allow for knowledge, reconstruction, or inference of the states of other minds? Answering that question involves the examination of the constraints and potentials of different means of signifying. To that end, we adopt a Peircean scheme, which makes a distinction between signals and signs. Peirce (1910/1967) specifies the concept of sign by distinguishing three types: *icons, indexes,* and *symbols.* Signals have a simple two-part structure. The "triggering" stimulus automatically releases patterned behavior. In contrast, signs represent by establishing different forms of relations between their three elements. Accordingly, all signs are triadic (three-part) structures that include reference to an object, a mental state (a representation of that object), and a sign vehicle (the medium of representation). However, the nature of the articulation of the three elements is different in each of the three types of signs.

Icons represent by virtue of physical similarity between sign vehicle and object and for this reason would not be sufficient for representing other minds. Indexes establish meaning through physical or temporal contiguity between the object and sign vehicle. Indexes may be sufficient to point to the functioning of a mental system behind observable behavior, but not precise enough to represent discernable mental states. Dennett's (1987) reading of a thermostat as an intentional system illustrates the imprecision of a "mental state index." According to Dennett, because a thermostat modulates output by processing input parameters, this indexes a central processor (a "mind"). It is impossible, however, to infer anything else about the nature of the mind on the basis of this analysis.

In contrast, the nature of the link between the referent and the vehicle in the symbol is arbitrary and conventional, and therefore does not require either physical resemblance or the natural relation that characterize the icon and the index, respectively.[3] Only in the symbol is there enough differentia-

---

[3]When we say that symbolic representation is arbitrary in nature, we do not mean to say that symbols are arbitrary in origin. Arbitrariness (or conventionality) arises as a potential limit in the continuum of "distancing" between referent and symbolic vehicle (where there is high differentiation between the material form of the vehicle and that which it is meant to represent).

tion between the material form of the sign and what is to be represented to allow for the representation of specific mental states.

Objective and Dual Subjective Modes of Analysis

The "abstract" or "objective" analysis of signs focuses on classifications and typologies of signs. This form of analysis can be supplemented by a functional study of signs-in-use. "Dual subjective analysis" focuses on subjects in inter- action and raises questions about processes and modes of coordinating sub- jectivities. Since, in this paper, understanding others' minds is conceptualized as a communicative phenomenon, it is precisely this form of analysis of signs- in-use that better suits our issue, the relation between types of signs and orders of intentionality.

One other piece of metatheoretical territory needs to be cleared before we can proceed: the place of historical analysis in the proposed framework. Conceptualizing human semiotic and intentional processes as modular, evo- lutionary discontinuities robs them of their history. They become fortunate accidents in the evolutionary history of the brain. However, as Vygotsky (fol- lowing Blonskii) once wrote, "To grasp in investigation the process of devel- opment of something in all of its phases and changes — from its emergence to its end of existence — is the essence of revealing its nature since 'it is only in movement that a body shows what it is'" (1960, p. 89). Following this line, it seems that another important feature to add to our framework is the consideration of the developmental history of symbolic processes. In general, history can be conceptualized through different units of analysis (the devel- opment of the function in the individual; the development of the function in the species) and therefore investigations of both phylogeny and ontogeny may lead to insights into symbolization and intentionality.

Flexibility and Change in Primate Vocal Communication

Basic forms of vocal communication date back at least 150 million years (Ploog, 1995). Each mode of communication is capable of providing certain types of information and is, simultaneously, limited in what can be captured or described. The nature of the information mediated by animal calls in gen- eral, and specifically by non-human primate calls, has long been a center of debate. In the past, non-human primate communication was thought of as fundamentally split from human symbolic processes. Non-human primate calls were conceptualized as mere emotional discharge with at most a signal or triggering function, with little or no self-awareness or consciousness of meaning implied. From this "big divide" perspective, it is difficult to think about the developmental continuity between calls and words despite the

evolutionary continuity commonly accepted between those who produce calls and those who produce words. Further research into primate vocalization has revealed a more complex picture, including forms of communication that go beyond emotional discharge and contagion, and may provide leads into their connection with symbolic processes (Cheney and Seyfarth, 1990; Ploog, 1995; Wrangham, 1977).

Ploog (1995) provides an example of emerging flexibility within signal-like behavior among non-human primates. This flexibility (in the marking of emotional states) is made necessary and facilitated by the complex social structures and hierarchies of certain primate groups. For example, Ploog describes the ontogenetic development of squirrel monkeys' system of emotional display. In early life, squirrel monkeys use a simple system of initially undifferentiated displays of emotional arousal. These displays are triggered or released by the presence of any conspecific (i.e., members of the same species). Over the course of ontogenesis, that system becomes differentiated and takes on new functions, such as influencing the behavior of other squirrel monkeys in agonistic, dominance, and courtship relationships. These displays can also occur in different forms depending upon the emotional state of the signaler and the social status of the intended receiver.

The appearance of flexibility within the initially signal-based system plays a key role in the dynamic structuring of the social group as well as the individual's social life within the group. Studies of other primate groups have illustrated the appearance of novel functions as a result of the complexification of systems of social display. For example, sonographic studies of vervet monkeys' alarm calls in the wild have indicated the genesis of referential function, which opens the possibility for the manipulation of the behavior of others. The structure and function of vervet alarm calls go beyond emotional release in that they (1) can be adjusted according to the audience (e.g., the lack of a call from a solitary monkey despite the presence of a predator [Cheney and Seyfarth, 1990]), (2) can be modulated according to environmental conditions (e.g., the abundance of food [Wrangham, 1977]), (3) can be delayed, absent, or variable in response to a given call, and, (4) do not spread by contagion. The significant modulation of the calls in relation to changing aspects of the physical and social environment speaks against the classification of those calls as signals. Such progressive complexification may also lead to the emergence of novel semiotic forms and functions — which, in turn, support the possibility of new psychological functions, including higher-order intentionality.

The Semiotics of Vervet Alarm Calls

It is our thesis that different communicative means (signal, icon, index, symbol) support different sorts of knowledge about other minds. Earlier in this paper, it was suggested that signals, for example, afford no information about other minds because of their automatic triggering. It was also suggested that only symbols suffice for the representation of other minds. We do not believe that vervet monkeys (upon which much of the literature focuses) are capable of representing the mental states of others. That would suggest, correctly, that we also do not believe that vervet communication utilizes symbolic signs. However, we do not want to fall into the usual trap of equating "non-symbolic" with "signal." The complexities presented by the primatological literature reviewed above suggest that another solution is possible.

The questions that follow are: (1) What do vervet calls refer to? (2) What is the quality of this reference? (3) How do the calls function differently for the different participants in communication? In our perspective, a productive way to proceed is to theorize possible intermediary forms that exist between signals and symbols. Ideally, such speculation would allow us to accomplish two goals simultaneously: (1) provide a clear semiotic grounding for observed forms of social behavior while not explaining away or granting too much and (2) help us understand the conditions for the emergence of symbols out of the restructuring of non-symbolic phenomena. Accomplishing the second goal is key for supporting our assertion that a single developmental process provides continuity among different semiotic forms and the psychological functions that depend upon those forms.[4] We will then illustrate *how* we suppose the same process of restructuring that underlies the emergence of symbolic functioning gives rise to the ability to represent other minds.

With that aim in mind, we suggest that vervet alarm calls function through *double indexical reference*. The term *double indexical* is intended to denote a sign form in which the vehicle holds an indexical relation (i.e., a habitual relation of contiguity) with two different referents.[5] In the case of the vervet alarm calls, we believe that both the external object and the caller's own internal, emotional states are indexed.

Specific support for the idea that these calls refer to an external object comes from laboratory studies in which vervets were habituated to another monkey's "intergroup *wrr*" (a type of call assumed to refer to the presence of another group — see Cheney and Seyfarth, 1990). In cases when the habitu-

---

[4]By stating this, we are highlighting the idea that continuity (the pre-existence and reorganization of elements) is always implied at a certain level of emergent (discontinuous) phenomena. Otherwise, the argument for emergence becomes synonymous with creationism.

[5]The prototypical Peircean index is established by a habitual relation of contiguity. Here we extend the definition to such a relation including more than one referent.

ated vervets were exposed to the same monkey's "intergroup *chutter*" (a type of call assumed to refer to the presence of another threatening group), habituation transferred. However, exposure to the same monkey's "leopard alarm call" produced dishabituation and the usual response of running into the nearest tree. This evidence suggests that monkeys, as well as being aware of the "identity" of the caller, compare different calls on the basis of some aspects of the external object to which they refer — and not on the basis of acoustic properties. In other words, calls that referred to identical things were treated the same even if they sounded different.

Vervets' alarm calls do not only indicate information about external referents; they also indicate the internal state of the caller. *For the caller*, vocalizations are both denotative and expressive (Cheney and Seyfarth, 1990, p. 168). This suggests that although the alarm call is given because of the presence of a predator, there is still an aspect of emotional expression inherent in the call. Although we understand these two referential components as hierarchically related (emotional expression being more basic than reference), the two functions are still analytically separable. Calls that require belief in the presence of some object embed the function of calls that are part of a more general/diffuse emotional reaction, which could include behavioral automatisms. To clarify, in the case of vervets, alarm calls often coincide with patterned behavior but that behavior is not automatic (it may or may not happen). This suggests that the vervets' emotional reaction has become a mental state: differentiated (but not separated) from the bodily reactions that are their most primitive manifestations. Support for this claim comes from Ploog (1995), who discusses evidence from ethological and neurological studies that indicate a relative dissociation of emotion-marking vocalizations and presumed motor correlates in primates. Specifically, Ploog has been able to demonstrate that the subcortical centers for vocalization can be selectively activated without the involvement of the motivational system.

The modulation of calls according to audience and circumstance further demonstrates the differentiation of the call from other components of the alarm response. For example, vervets sometimes emit a quieter alarm call when other vervets are not present. This suggests to us that the call *indexes* (and is not a necessary/automatic part of) the caller's mental state. In other words, we understand the indexical relation to involve some degree of differentiation between the call and the emotional state. In signals, the call is a necessary part of the general, diffuse emotional reaction; in an indexical relation, the call habitually co-occurs with, and eventually comes to represent, the mentalized emotional state.

Further evidence of the indexical nature of the referential links involved in vervet communication comes from the lack of contagion of emotional arousal among vervets. In a scenario involving simple contagion or mimicry,

the spread of emotional arousal occurs without mediation by anything but the arousal call itself. In this case, the call would function as a signal, triggering the full spectrum of emotional reactions (including more calls) in other monkeys of the group. However, we have already reviewed evidence that suggests this interpretation is not sufficient. Vervets seem to interpret each other's calls in terms of their external referents. Thus, alarm calls focus attention on the surrounding environment. Once the object being referred to is perceived, the road is open for the vervets to respond (or not) as the situation dictates.

On the other extreme of the semiotic scale, another alternative is that the spread of emotional arousal is mediated by a representation of the mental state of the other (i.e., symbolically). The evidence provided by Cheney and Seyfarth (1990) does not seem to support such a claim. Calls seem to be interpreted in terms of their external referent and probable behavior of the caller for the immediate future, but not in terms of the mental states of the caller. In other words, the fact that the listener may respond to the predator call in the same way as the caller responds to it does not imply that the listener has understood the caller's stance toward the referent. Rather, the listener, whose attention has been called to the referent by the alarm, responds independently, while taking situational factors into account. Indexical reference calls attention to the physical and temporal surround — the whole of which the index itself is habitually a part (Wallon, 1974, p. 221). Viewing the status of the call as indexical and, therefore, as "suggestion" (as opposed to the inflexible "command" nature of the signal) allows us to understand why any individual may display flexibility in its reactions to alarm calls; signals do not suffice and symbols are not necessary.

The Broken Triangle: Double Indexicals in Action

Having described the dual reference and indexical nature of any single vervet alarm, we can now approach our third issue: how those calls function and what sort of information about the other is made available within the communicative situation. It is here that the "dual subjective" mode of analysis becomes particularly useful. Although we may be able to say that one monkey *knows* what another is referring to when it hears the other give an alarm call, it does not follow that we can say that the listener understands the mental state of the caller — its stance toward whatever external object it is calling about.

Co-occurrence of references to an external object and to an emotional state does not imply the integration of those references.[6] That would lead to

---

[6]We use the term "integration" to refer to the process through which co-occurring parts are brought together in such a way that they constitute a new functional whole in which the components enter into a dynamically stable relation of mutual definition.

the simultaneous comprehension of the two references by the listener. Although any (vervet) listener can interpret a call in terms of its external referent, the internal referent is limited to the field of awareness of the caller. When communicating with the double indexical, the mental states of vervet monkeys are "private." This characterizes an *informational asymmetry* between caller and listener as to each other's mental states. The structure and functional properties of the double indexical are shown in Figure 1.
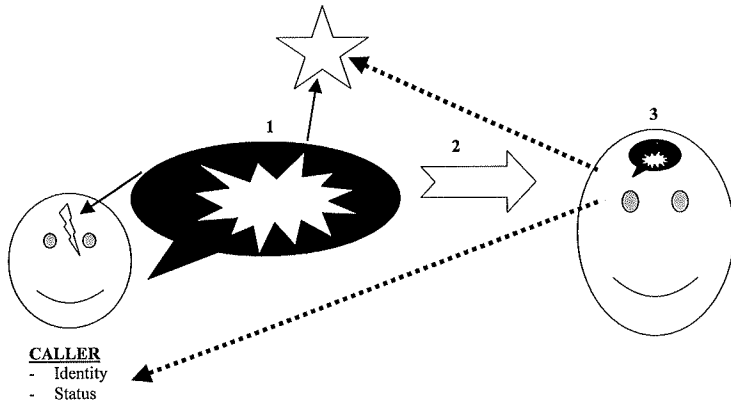

Figure 1: Schematic depiction of the asymmetry in information available to the caller and the listener. Here, the first monkey sees an object (star) and experiences some emotional state (represented by a lightning bolt). Its call indexes both the object and the emotional state. (1) The call is heard by the listener (2) who interprets the call in terms of an external referent of a certain general class object and the caller's identity and status (3) but not the caller's emotional states.

It is this kind of representational link, and the participants' asymmetric access to information provided by the link, that allows the deceptive use of alarm calls (i.e., a call in the absence of a predator to influence the behavior of others) without the representation of others' mental states.

The Semiotics of Deception Among Vervets

Although it is tempting to explain the way deceptive calls function in terms of higher-order intentionality (including a representation of and attempt to manipulate the mental states of others), the inconsistent behavior exhibited by the calling monkeys makes this seem implausible. For example, one vervet, Kitui, gave a leopard alarm call in order to keep a new male from migrating into the group (Cheney and Seyfarth, 1990). After giving this call and watching the intruding male run up the nearest tree, Kitui himself came down from his perch, walked across the field (where the supposed leopard should be), and climbed into a tree near the intruding male. Had Kitui

intended to influence the *beliefs* of the intruder, he would not have exhibited behavior that could be interpreted by the intruder as contradictory.[7]

In this example, deception is accomplished because the alarm call is interpreted (by the listener) in terms of an external referent and *not* the intentional states of the caller. In this way, the double indexical suffices as a tool for the manipulation of others' behaviors, but does not "afford" the manipulation or inference of others' thoughts and beliefs through verbal communication. This is a mechanism through which deception is made possible without such representation. Whereas deception is most frequently conceptualized as involving representation of others' intentional states (and their purposeful manipulation), we agree with Cheney and Seyfarth (1990) in their assessment of vervets' deception as a phenomenon that can be explained through a more modest account involving only first-order intentionality. However, rather than giving this claim only a logical grounding, we hope to have added a semiotic account showing that the double indexical allows for deception without representation of others' mental states. In other words, the caller acts intentionally, but effectiveness of its action does not depend on the listener's interpretation of the caller's intentional states (thoughts, beliefs, desires, etc.).[8]

It is interesting to note that early versions of human deception are just as superficial as Kitui's attempt. Cheney and Seyfarth (1990) give the example of a three-year old child who denies having been to the cookie jar despite crumbs on its face. Humans, though, are able to go beyond this form of deception. For example, humans do things like faking a cough when calling into work sick to try to avoid having their pay docked. This kind of attempt to manipulate involves at least third-order intentionality: the caller wants the other to believe that she believes that she is sick.

---

[7] If beliefs were at stake here, Kitui's actions following the deceptive call would have nullified the effectiveness of his ploy. If the other monkey was capable of interpreting Kitui's actions in terms of beliefs about the presence of the leopard, the intruding male could have easily concluded that there was, in fact, no attacking leopard and continued his pursuit of the group unworried. This, however, was not the case (cf. Cheney and Seyfarth, 1990). However, as one reviewer suggested, there will always be a plethora of possible interpretations of any action. This is certainly true — even in the case of human beings — and it is a general metatheoretical and methodological issue for the social sciences. More specifically, Povinelli and Vonk (2003) argue convincingly for the point that no experiment in which theory of mind coding derives from behavioral abstraction will ever suffice to ultimately constrain interpretation. The interpretation of Kitui's actions that we offer here stays close to Cheney and Seyfarth's, as does our claim that higher-order intentionality is not necessary in order to explain those actions. While Cheney and Seyfarth come to that conclusion through a preponderance of evidence regarding non-human primate communication in general and the specifics of this example in particular (e.g., the difference in social status between Kitui and the intruder), our goal in this section of the paper is to provide an explanation, based upon our concept of the "double indexical," of *exactly how* this kind of deception could possibly work without having to appeal to higher-order intentionality.

[8] It is one of the main hopes of this paper that links can be made between types of signs used in communication, orders of intentionality, as well as informational and behavioral possibilities.

Trickery like this involves a representation of — and deliberate attempts to modify — the mental states of the other and therefore requires symbolic media. Symbols are required because they perform their referential function irrespective of the need for usual co-occurrence (indexes) or likeness (icons).[9] In the case of the phone call, the message is largely conveyed with the help of symbolic elements that are used for many other, nonspecific purposes and may thus serve in the communication of an unlimited number of messages. To summarize, there is one basic difference between the deceptive scenarios described. In the vervet case, double indexical reference is used to manipulate the behavior of others. In the human case, symbolic reference is used to alter the beliefs of others.

The Semiotics of Non-deception: Perspectival Description

The symbolic interpretation of the intentional states of others is even part of communication in non-deceptive scenarios. In contrast to indexes and icons, which allow for some degree of knowledge about the referent regardless of knowledge about the addressor, symbols provide information about both the world and the addressor. For example, if Maria tells Cecilia that the beach Maria visited was beautiful, Cecilia constructs knowledge about both the location and Maria's relation to it. Symbols are "perspectival" (Tomasello, 1999) and a medium adequate for the development of second-order intentionality. In contrast, double indexicals allow reference to an external referent, but not an understanding of the caller's stance. From our semiotic perspective, this implies a linking of the internal and external referents which were disjunct in the double indexical, creating the familiar triadic form of the fully integrated symbol.

Having established the affordances of the two forms, we can speculate about the genetic relations between these two. In other words, can continuity be posited between the double indexical nature of the vervet calls and the symbolic reference in human words? How might that development have allowed for emergence of the ability to infer the mental states of other beings (conspecifics or not)?

---

[9]Otherwise, communication would have to take place through the presentation of symptoms that index illness — similarity to previous instances of illness or other nonsymbolic evidence. Although such "symptoms" (e.g., a — potentially fake — cough or achy voice) are often presented while "calling in sick," those presentations are made only to supplement or make the symbolic presentation more believable.

Symbols in the Making

As we have discussed, the ability to attribute mental states and beliefs to others appears to be absent in monkeys due to lack of integration in the double reference of their calls. Our hypothesis is that, not coincidentally, this new function is brought about by the same structural change that transforms the double indexical into a symbol. In this sense, we understand the double indexical as a genetically intermediate form between simple indexes, in which reference is possible through co-occurrence or pairing, and symbolic reference, which must include the integration of the two links. We can approach the problem of the genesis of semiotic forms by analyzing some examples, after which more general principles can be elaborated.[10]

One instance of the development of sign forms involves the move from signals to the double indexical. This move implies steps of differentiation and entails the emergence of first-order intentionality. In the signal, there is no consciousness or mediation involved — reference and stance are not yet differentiated into identifiable components. Specific external objects elicit a signal and the behavior of which that signal is a part spreads through contagion (is self-spreading). Thus, the signal relation has two poles. On one hand, the signal is open to the "triggering" function of the environment. On the other hand, the signal functions to "release" a pattern of behavior.

This "fit" between an environmental stimulus and the reaction can be thought of as an elementary form of a relation to the object and a relation to the internal state. For instance, one can imagine a simple system of signals, which function in the food domain: one pattern of behavior that occurs when the opportunity to eat food arises and one pattern of behavior that occurs when the danger of becoming food appears. Each of these reflex-sets involves a general sort of object (one that can be eaten or one that can eat) and a coincidental bodily-emotional state (which can be glossed as happiness or distress). None of these categories is explicitly conceptualized, but is implicit in the fit between the external environment and different features of the particular animal's structured system of potential responses.

In signal form, the three components (vocalization, referent, stance) are undifferentiated. In the double indexical, however, those components become differentiated and the vocalization takes on representational qualities vis-à-vis the referent and internal stance. It is important to note that these three components (vehicle/call, reference, and stance) are the "raw materials" for symbolic reference.

---

[10]In this analysis, we acknowledge the influence of Werner and Kaplan's (1963) *Symbol Formation.*

Continuing this developmental pathway, the move from double indexicals to symbols implies the integration of the indexicals into a single system that functions in a new way. The integration of the two indexicals is a structural change in which the previously existing three elements become synthesized into a new functional whole, transforming the character of each. This integration is made clearer by the fact that the character of each element can only be defined in relation to the others.
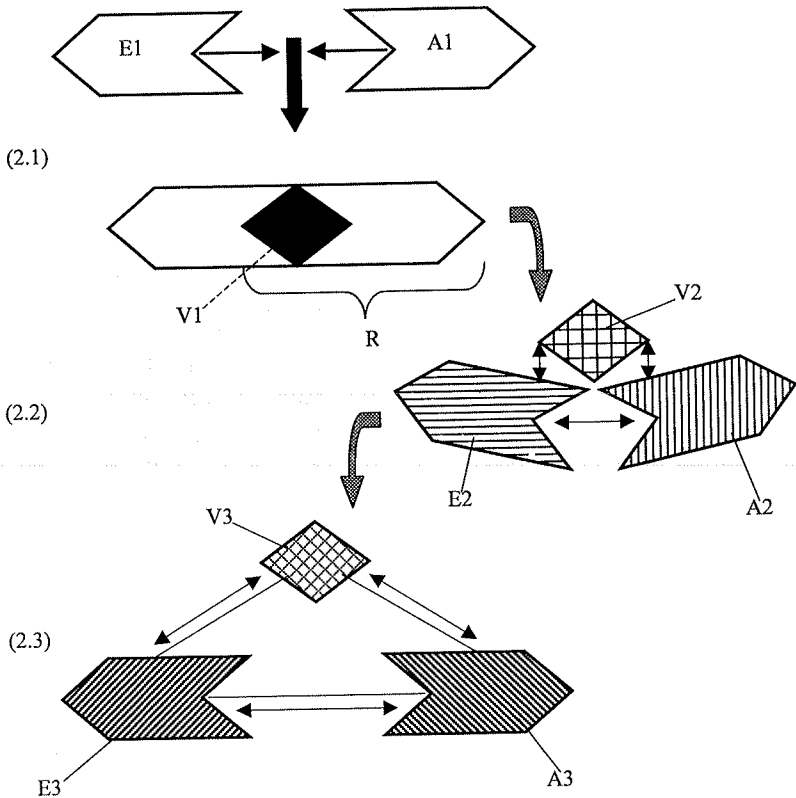

Figure 2: Genetic progression from signals to symbols. In a signal (2.1), an environmental stimulus (E) comes within the range of awareness (A) of the subject, leading to the release of an undifferentiated alarm response (R) the shaded part of which represents its vocal aspect (V1). In the double indexical (2.2), all elements undergo differentiation and distancing (represented by double-headed arrows). The vocal aspect is distanced (V2) and comes to refer indexically both to the caller's internal emotional state (A2) and to some specific aspect of the environment (E2). In 2.3, the symbol is represented as a highly distanced structure in which the vertices, vocalization (V3), conceptual interpretant (A3), and referent (E3) each acquire their functional properties as a result of a process of differentiation and mutual definition of the previously established elements.

In the move from double indexical to symbol, the vocalization becomes less contiguous with the referent and can perform its representational function without a history of co-occurrence. The vehicle becomes denaturalized in that there is less and less that links it to the referent, except convention. However, it is the abstract nature of the relation between symbolic vehicle and referent that gives the symbol its depth and flexibility as a representational tool.

The links to the external referent and to the stance enter a process of mutual definition. What was previously an ill-defined, diffuse emotional reaction becomes more highly specified and conceptual. The dynamics of specification can be best understood if we imagine the diffuse emotional reaction becoming distanced from its referents and becoming differentiated into specific stances that are oriented toward different aspects of those referents.[11] This involves a qualitative transformation of emotion into a conceptual interpretant of that particular object. On the other hand, the emotional stance highlights certain aspects of the percept (transforming the percept into an interpretant) and relates it to other interpretants. It is in this process of mutual definition that the emotional stance is transformed into meaningful and interrelated concepts. With this new mediational tool, knowledge about the object as a circumscribed entity becomes possible. In sum, we understand the qualitative transformation of signs to take place according to the general principles of differentiation and integration.

Organic and Synthetic Symbols

It is useful at this point to contrast this account of natural semiogenesis with the comparatively artificial-looking process of symbol acquisition observed in ape language projects. In those projects (Rumbaugh, 1977; Savage–Rumbaugh, 1986) chimpanzees learn to pair lexigrams to objects. It is important to note that, in these pairings, the sign is externally imposed. The issue is that the sign is not formed through a process that incorporates the stance of the animal toward the object (the lexigram for apple is paired to apple regardless of the chimp's stance). In these laboratory pairings, stances, although they surely exist, are alienated from the system (cf. Rumbaugh, 1977). The nature of these laboratory pairings seems to hinder the chimps' development (rather than acquisition) of a fully functional symbolic system. More promising (from an admittedly symbol-centered perspective) seem to

---

[11]This is represented in Figure 2.2, where the vocalization takes on a cross-hatched pattern. The horizontal component of the cross-hatching corresponds to the horizontal striping of E2, which is the specific aspect of the environment to which the vocalization refers.

be the instances where apes combine or modify previously known signs to form new "words."[12]

In contrast, the symbol that emerges from our genetic account is one in which the interpretant can be traced back to an undifferentiated emotional stance. The suggestion that originally undifferentiated emotional phenomena are the "precursors" of concepts should not by taken to mean that the individual, affective link to a referent is altogether lost. Rather, concepts retain their emotional grounding and the two levels of complexity (emotional and conceptual) never lose touch. This allows for a sign vehicle that is highly particular and personal but distanced and conventional enough to allow others insight into the meaning of the sign.

Sociogenesis of Symbolic Function

It is the human capacity to integrate reference to an external object with reference to internal states that paves the way for the development of the symbolic function. This integration is not an isolated act of the individual. It is the social nature of human groups that provides and supports the child's integration of these two spheres of reference.

It may be useful to draw attention back to the notion of the informational asymmetry within the primate communicative situation because a different kind of asymmetry may be at the heart of the development of symbolic reference during human ontogeny. It has been argued that, in the vervet vocalizations, the listener interprets calls in terms of their external referents and not in terms of the internal stance of the caller. Therefore, the listener signifies the call *less* than the caller. For the caller, the vocalization carries meanings and performs functions that are not available to the listener. Human infant–adult communication, likewise, exhibits asymmetry, but in another direction. When the infant vocalizes, social others potentially signify *more* than the child does.[13] That is, caretakers may interpret meanings

---

[12]See Menendez and Patterson (1994) on Koko's altering established hand signs and the establishment of new referents for these new configurations. In one instance, for example, by altering the hand sign for lettuce, a sign is established for "browse," a more general category of leafy greens.

[13]This sort of over-interpretation cannot be reduced to a one-way or automatic "imposition" of meaning. The interpreter's process of hyper-attribution of meaning is as important as the producer's orientation to the other or "openness" to social support, which is facilitated by the child's and the adult's emotional interchange, which precedes all other forms of communication. Within this "primordial shared situation," there exists the least possible differentiation between self, other, and external object and interaction occurs in sensory–motor–affective terms (Werner and Kaplan, 1963, p. 42). This allows the infant to share with the other a relation to sources of pleasure, apprehension, excitement, distress, rather than experiencing the world in isolation.

into babies' cries that, in effect, tell the baby what his or her own cry "really" means.[14]

The gradual appearance of coordinated patterns of acting and reacting underlies much of early human communicative development (Lyra and Winegar, 1997). Human babies are born into a hyper-social and hyper-symbolic world in which others interpret babies' actions, cries, and vocalizations even before the babies themselves can self-consciously use those actions to convey meanings or ideas.[15] Referential functions of the child's early utterances are established as the child actively connects the utterance, the world and the social reactions to the utterances. At first, reference is only to undifferentiated referents and happens in highly personal, contextualized, and

---

[14]In fact, human hyper-attribution of meaning seems to be a phenomenon that is not restricted to our interactions with infants of the same species. A dramatic example of interspecies hyper-attribution of meaning can be observed in the interpretation of gorilla signing. Consider the following example from Koko's internet chat (Gordon, 1998). In the interaction, MiniKitty (MK) is a person online, chatting with Koko (K) through Penny (P), one of Koko's human "parents," who simultaneously reads and signs the incoming messages to Koko and (importantly) interprets her responses.

> MK: Koko, are you going to have a baby in the future?
> P: Koko, are you going to have a baby in the future?
> K: Koko — love eat . . . sip.
> MK: Me too!
> P: What about a baby? Are you going to have a baby? She's just thinking . . . and her hands are together . . . .
> K: (*signs*) Unattention.
> P: Oh, poor sweetheart! She said, "Unattention." And what that is, she covers her face with her hands . . . which means it's not happening, basically, or it hasn't happened yet . . . . I don't see it.
> MK: That's sad.
> P: She's responding to the question. In other words, she hasn't had one yet, and she doesn't see a future here. The way the situation is actually with Koko and Ndume [the male gorilla], she has two males to one female, which is the reverse of what she needs. I think that is why she said that, because in our current situation, it isn't possible for her to have a baby. She needs several females and one male to have a family. (Gordon, 1998, p. 12, italics added)

Although the sign-word "Unattention" may have been established for Koko through sign-instance pairings and therefore functions indexically, Penny (over-)interprets Koko's sign in a way that transcends contingency (i.e., symbolically). By highlighting this hyper-attribution of meaning, we wish to draw attention to the pervasiveness of this phenomenon and to speculate on the function of this apparently irrational behavior.

[15]The classical analysis of the development of pointing — stemming from Wundt and known largely through the writing of Vygotsky (1960) — illustrates that (1) social others interpret the child's unsuccessful grasping effort as an indication of both some external object and the child's own desire or orientation, and (2) the child actively connects her own actions and the others' reactions, allowing the reaching action to become an indicator for the child herself. In this interpretation, the child is the *last* to become conscious of the meaning of her act. In this process, the child gradually overcomes the asymmetry and thereby acquires the ability to use gestures (and later on, other modes of representation) as signs intentionally within a communicative situation.

emotional ways. This point has a history in the literature on communicative development. Guillaume (1927) uses the term "monoremes" to refer to the first utterances as predicates of the situation, global and concrete in nature; Werner and Kaplan (1963), following Gregoire (1937–1947, cited in Werner and Kaplan, 1963), define monoremes as one-unit referential vocalizations that appear prior to the co-emergence of words and sentences but nevertheless refer to whole situations. Finally, in her discussion of early cognition, Nelson (1996) points out that single-unit utterances take on a pragmatic, rather than symbolic, function and are used to mark the child's desired experience of instantiations of social routines or "events."

It is worth emphasizing the idea that these fuzzy references are not "cold" and categorical, but indicative of the child's hedonic states or emotional relation to the events. Thus, the child's first words can be seen as having a similar structure to that of the double indexical primate monkey calls. They refer independently to an emotional state that co-occurs with an object or event, and also (imprecisely) refer to that fuzzy object or event. In the human case, the adult's interpretation of and reactions to the child's first communicative efforts provides a model or "pull" which facilitates the child's own internalization of those meanings and progression toward symbolic communication.

In turn, the child establishes, through early affective relations, an orientation toward (or openness to) the caretaker. That orientation makes the caretaker's interpretations an especially powerful form of support. That is, social others provide external support that allows a developing child to perform in ways that it never could by itself. Such support also acts as a catalyst for the child's development of her own abilities. In this case, social others, while engaged in communicative efforts with the child, provide the link between the child's separate indexes, allowing those more primitive signs to function — first for others, then for the child herself — as true symbols (i.e., as triadic structures that integrate vehicle, referent, and representation or interpretant).

Conclusions

We opened this paper by highlighting a basic issue: continuity and discontinuity in the co-evolution of semiotic and intentional processes. There can be no doubt that there is a "difference" between human and non-human primates, but that difference does not imply a sudden evolutionary discontinuity. No magical force granted a lucky evolutionary ancestor a functioning theory of mind module; nor did that force grant the capacity to represent symbolically. Those are novel functions that appeared as part of a continuous process of development (differentiation and integration) that has occurred over evolutionary (phylogenetic) time and continues to happen, everyday, on the ontogenetic scale. By taking a structural stance toward signs and their

development, we hope to have demonstrated that the raw materials necessary for the emergence of symbolic representation are present in both the child and the non-human primate. That is not to say that we propose a recapitulationist theory, although we are unwilling (and perhaps unable) to offer an evolutionary account of the emergence of symbols.

We believe that the integration of the double indexical is the step needed to pass through the semiotic threshold into symbolic functioning and that this cannot happen in the solitary mind of an evolutionary prodigy. Our theoretical exercise here only offers *an* answer to the question of what might have happened over the course of phylogeny to give rise to symbolic capacities, through which all higher mental processes emerged. Through our description and developmental analysis of different types of reference, we hope to have shown that some of the elements used to characterize the human–nonhuman difference ("theory of mind") are epiphenomena of a more fundamental semiotic innovation — an innovation which, like any other, does not deny its own history, but rather re-writes it — and one which also awaits its own obsolescence.

References

Baldwin, J.M. (1894). Imitation: A chapter in the natural history of consciousness. *Mind, 3,* 26–55.

Baron–Cohen, S. (1995). *Mindblindness.* Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Bartsch, K., and Wellman, H.M. (1995). *Children talk about the mind.* New York: Oxford University Press.

Boesch, C. (1993). Aspects of transmission of tool use in wild chimpanzees. In K.R. Gibson and T. Ingold (Eds.), *Tools, language, and cognitive evolution* (pp. 171–183). Cambridge, England: Cambridge University Press.

Carruthers, P., and Smith, P.K. (Eds.). (1996). *Theories of theories of mind.* Cambridge, England: Cambridge University Press.

Cheney, D., and Seyfarth, R. (1990). *How monkeys see the world.* Chicago: University of Chicago Press.

de Lima, S.G., and Surgan, S. (in press). Innovation in imitation: Relating opposites in Tomasello's theory of cultural development. *Clark working papers on developmental psychology, 3.* Worcester, Massachusetts: Clark University Press.

Dennett, D.C. (1971). Intentional systems. *Journal of Philosophy, 68,* 68–87.

Dennett, D.C. (1987). *The intentional stance.* Cambridge, Massachusetts: MIT Press.

Dennett, D.C. (1996). *Kinds of minds: Toward an understanding of consciousness.* New York: Basic Books.

Dennett, D.C., and Haugland, J. (1987). In R. L. Gregory (Ed.), *The Oxford companion to the mind* (pp. 383–386). Oxford, England: Oxford University Press.

Fogel, A., and Branco, A.U. (1997). Metacommunication as a source of indeterminism in relationship development. In A. Fogel, M.C.D.P. Lyra, and J. Valsiner (Eds.), *Dynamics and indeterminism in developmental and social processes* (pp. 65–92). Mahwah, New Jersey: Erlbaum.

Gordon, W. (1998). Koko's internet chat. *Gorilla: Journal of the Gorilla Foundation, 21,* 11–12.

Grégoire, A. (1937–1947). *L'apprentissage du langage.* Paris: Droz.

Guillaume, P. (1927). Le développement des éléments formels dans le langage de l'enfant. *Journal de Psychologie, 24,* 203–229.

Huffman, M.A. (1984). Stone play of *Macaca fuscata* in Arashiyama B troop: Transmission of non-adaptive behavior. *Journal of Human Evolution, 13,* 725–735.

Itani, J. (1958). On the acquisition and propagation of a new food habit in the natural group of the Japanese monkey at Takasaki-Yama. *Primates, 1,* 84–86.

Kandel, E.R., Schwartz, J.H., and Jessell, T.M. (Eds.). (1995). *Essentials of neural science and behavior.* Norwalk, Connecticut: Appleton and Lange.

Kawai, M., Tsumori, A., and Motoyoshi, I. (1963). A confirmatory analysis of the new acquired behaviors of the Japanese monkeys on Koshima Island. *Primates, 4,* 115–116.

Kawamura, S. (1959). The process of subculture propagation among Japanese macaques. *Primates, 2,* 43–54.

Kummer, H. (1971). *Primate societies: Group techniques of ecological adaptation.* Chicago: Aldine.

Legerstee, M. (1997). Changes in social–conceptual development: Domain specific structures, self-organization, and indeterminism. In A. Fogel, M.C.D.P. Lyra, and J. Valsiner (Eds.), *Dynamics and indeterminism in developmental and social processes* (pp. 245–260). Mahwah, New Jersey: Erlbaum.

Lyra, M.C.D.P., and Winegar, L.T. (1997). Processual dynamics of interaction through time: Adult–child interactions and process of development. In A. Fogel, M.C.D.P. Lyra, and J. Valsiner (Eds.), *Dynamics and indeterminism in developmental and social processes* (pp. 93–109). Mahwah, New Jersey: Erlbaum.

Mandler, J.M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99,* 587–604.

McGrew, W.C. (1992). *Chimpanzee material culture.* Cambridge, England: Cambridge University Press.

Meltzoff, A., and Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. In S. Baron–Cohen, H. Tager–Flusberg and D.J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 335–366). Oxford, England: Oxford University Press.

Menendez, D., and Patterson, F.G.P. (1994). Innovative signing. *Gorilla: Journal of the Gorilla Foundation, 18,* 2–3.

Mitchell, D.E. (1980). The influence of early visual experience on visual perception. In C.S. Harris, (Ed.), *Visual coding and adaptability* (pp. 469–484). Hillsdale, New Jersey: Erlbaum.

Nelson, K. (1996). *Language and cognitive development: Emergence of the mediated mind.* Cambridge, England: Cambridge University Press.

Nishida, T. (1986). Local tradition and cultural transmission. In B.B. Smuts, D.L. Cheney, R.M. Seyfarth, R.W. Wrangham, and T.T. Struhsaker (Eds.), *Primate societies* (pp. 462–474). Chicago: University of Chicago Press.

Peirce, C.S. (1967). Manuscript 654.7. *Manuscripts.* Cambridge, Massachusetts: Harvard University. (Originally published 1910)

Ploog, D.W. (1995). Mutuality and dialogue in non-human primate communication. In I. Markova, C. Graumann, and K. Foppa (Eds.), *Mutualities in dialogue* (pp. 27–57). Cambridge, England: Cambridge University Press.

Povinelli, D.J. (1996). Chimpanzee theory of mind?: The long road to strong inference. In P. Carruthers and P.K. Smith (Eds.), *Theories of theories of mind* (pp. 293–329). New York: Cambridge University Press.

Povinelli, D.J., and Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognition Sciences, 7,* 157–160.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1,* 515–526.

Rumbaugh, D.M. (1977). *Language learning by a chimpanzee.* New York: Academic Press.

Savage–Rumbaugh, E.S. (1986). *Ape language: From conditioned response to symbol.* New York: Columbia University Press.

Tomasello, M. (1994). The question of chimpanzee culture. In R.W. Wrangham, W.C. McGrew, F.B.M. DeWaal, and P. Heltney (Eds.), *Chimpanzee cultures* (pp. 301–317). Cambridge, Massachusetts: Harvard University Press.

Tomasello, M. (1998). Social cognition and the evolution of culture. J. Langer and M. Killen (Eds.), *Piaget, evolution and development* (pp. 221–246). Mahwah, New Jersey: Erlbaum.

Tomasello, M. (1999). *The cultural origins of human cognition: An essay*. Cambridge, Massachusetts: Harvard University Press.

Tomasello, M., Kruger, A., and Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495–511.

Valsiner, J. (1997). *Culture and the development of children's action*. Chichester, England: Wiley.

Vygotsky, L.S. (1960). *The development of higher psychical functions*. Moscow: APN.

Wallon, H. (1974). Las relaciones de lo significante a lo significado. In H. Wallon, *Del acto al pensamiento* (pp. 151–168). Buenos Aires: Editorial Psique.

Wallon, H. (1975). Las emociones en el comportamiento humano. *Las origines del carácter en el niño* (pp. 81–95). Buenos Aires: Nueva Visíon.

Werner, H., and Kaplan, B. (1963). *Symbol formation: An organismic–developmental approach to the psychology of language*. Hillsdale, New Jersey: Erlbaum.

Wrangham, W.R. (1977). Feeding behaviour of chimpanzees in Gombe National Park, Tanzania. In T.H. Clutton–Brock (Ed.), *Primate ecology: Studies of feeding and ranging in lemurs, monkeys, and apes* (pp. 503–538). New York: Academic Press.