

wife, who has a long experience with my reactions to “the self-serving, right-wing politics of the rich,” feels my whole body tense up as I start to rise to the bait. She is uneasy about what is about to occur, and she places her hand very lightly on my hand, where I have gripped the table, ready to pull myself forward in my seat. The warmth and intimacy of her gesture is utterly inconsistent with the publicness and tenseness of the social context in which it occurs. My attention is sharply refocused on her, on the human (rather than the political) dimensions of the conversation, and, ultimately, on the humor and absurdity of the situation.

But let us now suppose that in addition to placing her hand on my hand, she had leaned over to whisper in my ear, “Easy does it!” It is still possible to accommodate this new move – the speech move – within the dance metaphor, speech being just another of the many modalities that are tuning my behavior, instant by instant, as the dinner party goes forward. But it is not easy to avoid the feeling that *this* dance move is special. True, the intimacy of a whisper, the softness of the tone with which it is delivered, even the *tickle* of it, all are akin to the intimacy of her hand placed on mine. But in whatever way these features have meaning, the words “Easy does it!” seem to have meaning in a further sense. It is as if your partner on the dance floor, wishing the two of you to swing about, instead of leaning into your body and throwing you ever so slightly off balance, had spoken the command, “Let’s swing!”

The difference does *not* seem to be in the tool relation between the two peractants. If my momentary goal in the dance is to twirl my partner glamorously under my arm, then her twirling is a tool in my grand design. If her momentary goal is for us to swing around together in an exalting display of our “coupleness,” then my swinging is a tool in *her* grand design. Thus, for a dancing couple, the place where each grips the other is the *handle* of a tool. What makes a spoken command a different sort of move in the dance is that the handle in that case is a *sign*, instead of the actual grip on the partner’s body.

Nobody would deny that all chimpanzees are adept at using social partners as tools. What seems to set the ape language of research chimpanzees apart from that of their wild conspecifics is their use of signs as handles. We heartily agree with the authors that the dance metaphor does more than the information-transfer metaphor to help us understand these animals’ laboriously acquired talent at sign use. Even more useful would be an elaborated version of their dance metaphor in which the dance is a square dance and every participant is both caller and dancer.

## Does the use of the dynamic system approach really help fill in the gap between human and nonhuman primate language?

Jacques Vauclair

Centre for Research in Psychology of Cognition, Language and Emotion,  
University of Provence, 13621 Aix-en-Provence Cedex 1, France.  
vauclair@up.univ-aix.fr <http://www.up.univ-mrs.fr/wpsycle/>

**Abstract:** The highly recommended transposition of the dynamic system approach for tackling the question of apes’ linguistic abilities has clearly not led to a demonstration that these primates have acquired language. Fundamental differences related to functional modalities – namely, use of the declarative and the form of engagement between mother and infant – can be observed in the way humans and apes use their communicatory systems.

I can only be in sympathy with Shanker & King’s (S&K’s) proposal that dynamic systems paradigms are better suited than information processing models for describing and explaining language achievements in any species. In effect the kind of approach advocated in the target article places the investigation of linguistic abilities in its proper context, namely, that language and related behaviors are worked out in humans through a dynamic system of

coregulations and coconstructions of cognition between social partners. In that respect, the paradigms on which S&K rely in their propositions are reminiscent of the Vygostkian approach to human cognition and language (e.g., Vygotsky 1962), also known as the “distribution cognition” approach (Johnson 2001).

Two remarks are in order before discussing the use and output of the dynamic system approach in ape language research (ALR) studies. Although S&K do an excellent and fair job of summarizing ALR studies, two aspects of potential importance for the paradigm they defend are lacking in their review. First, the comparative literature provides evidence that marine mammals possess abilities of a complexity comparable to that reported for chimpanzees and bonobos. The California sea lions trained in a symbolic matching technique using gestural signs (e.g., Gisiner & Schusterman 1992), could comprehend at least 20 signs (modifiers, objects, and actions) in about 200 three-sign combinations. Experiments with dolphins that followed the same basic experimental procedure also demonstrated excellent abilities of these animals in comprehending imperative as well as interrogative strings (e.g., Herman et al. 1993). Interestingly, the studies with marine mammals were carried out without apparently following the dynamic approach, for those studies were based on the more traditional information-processing models.

Second, the achievements concerning language-trained apes are characterized by a gap between the comprehension of the communicative signals they have been exposed to and their abilities to produce strings of gestures or lexigrams or both. For example, the bonobo Kanzi can comprehend quite complex instructions given to him in English (Savage-Rumbaugh et al. 1993), but his productions of gestures and lexigrams are rather poor compared to his comprehension. Thus, the most complex of Kanzi’s “word” combinations involves three items (usually made of two gestures and one lexigram; Savage-Rumbaugh et al. 1986). In humans, lexical comprehension and production work in parallel (e.g., Harris et al. 1995; Werker & Tees 1999). The gap observed in the bonobo makes it difficult to interpret the ape’s achievements in evolutionary terms unless we assume that comprehension has evolved in the absence of language, in which case we may question the adaptive function of comprehension in the absence of production. These issues raise the problem of the role of the human-defined environment (through a process described as enculturation; Tomasello et al. 1993) on the emergence of the cognitive and communicative skills involved in language that may have no equivalent in spontaneous behaviors and on which no selective pressure has supposedly acted (Vauclair 1996).

The main point I want to make bears on the linguistic abilities reported for the bonobo Kanzi, which are undoubtedly, as S&K correctly state, the most impressive among those exhibited by any nonhuman species. Considering that the primary function of language concerns the exchange of information about the world, and following the pioneering work of Bühler (1934), two principal functions (or modalities) can be distinguished in the linguistic as well as prelinguistic communication among humans (Bates 1979). The first function is the *informative function* and it takes two forms: a declarative form that represents states of the world (e.g., “Jim comes”) and an interrogative form. The second function is *injunctive* (imperative) and *exclamatory* and mostly expresses itself with requests and demands (e.g., “Come!”). Developmental work with young children has shown that the use of declaratives (e.g., Bassano & Maillochon 1994; Wetherby et al. 1988) becomes the dominant mode of communication between 1 and 2 years of age (about 60% of all utterances). Declaratives (Bates et al. 1975) can be words or gestures, and they function not primarily to obtain a result in the physical world but to direct another individual’s attention (its mental state) to an object or event, as an end in itself.

It is quite clear from the published work that the imperative function appears to be the main (if not exclusive) mode used by the “linguistically” trained apes. In the case of Kanzi, studies reveal (Savage-Rumbaugh et al. 1985; 1986) that, unlike human chil-

dren who use language to make indicative or declarative statements, about 96 percent of Kanzi's productions are requests. Thus, the difference between Kanzi's modalities of communication and the typical declarative mode observed by humans is striking. In effect, communication in the apes has essentially an imperative function (this appears indeed to be the rule for all animal species and this mode is sufficient to fulfill the biological requirements of, for example, warning again predators). In contrast, humans use not only linguistic signs but also prelinguistic communication means such as gestures (e.g., pointing) for both imperative and declarative purposes, as when two persons share an interest toward a third person (Vauclair 1996; in press). This is further illustrated in a study in which we assessed the behavior of human-human, human-ape, and ape-ape pairs during object manipulation (Bard & Vauclair 1984; Vauclair 1984). One of the pairs we observed involved the bonobo Kanzi, his adoptive mother Matata, and a human caretaker (Kanzi was 8 months old when the observations started). We found that the communicative styles expressed by the adult toward the infant differed between species; thus, in contrast to the human adults, the ape adults rarely acted on the objects in ways that would direct or engage the infants' attention. Because our study was devised to focus on the factors of coattention and coregulation between infants and competent partners, it fit exactly with the principles of the dynamic system paradigm or distributed cognition (Johnson 2001). Our results clearly suggested that the patterns (e.g., joint attention to objects) which characterize mother-infant communication in humans were missing in the chimpanzees and the bonobos examined within our comparative framework.

## Human expression and experience: What does it mean to have language?

Yves-Marie Visetti<sup>a</sup> and Victor Rosenthal<sup>b</sup>

<sup>a</sup>LATTICE-CNRS-ENS, 92120 Montrouge, France; <sup>b</sup>INSERM XR-324, Centre Paul Broca, 75014 Paris, France. yves-marie.visetti@ens.fr victorro@broca.inserm.fr

**Abstract:** We support Shanker & King's (S&K's) proposal for a dynamic systems approach in ape language research, but question their vision of what it means to have language. Language plays an essential role in the making of the human mind. It underlies any kind of human interaction and codetermines perception and action. Moreover, what gives human thought the very characteristic architecture of textuality criterially requires a third party.

We are sympathetic with Shanker & King's (S&K's) plea for a dynamic systems approach in ape language research and share their critique of the information processing model. But we find their proposal too general and unspecific with respect to communicative abilities of humans and apes, and we question their vision of what it means to have language. In our view, the dance metaphor may be an illustration of interactional synchrony in a dynamic system, but it captures little of what may characterize human communication, and we argue the insufficiency of a dyadic relationship for the emergence of language skills in favor of a more complex triadic structure. What S&K call a "new paradigm" is reminiscent, both with respect to the idea of dynamic systems and with respect to the principles of ape cognition research, of the Gestalt paradigm formulated by Wolfgang Köhler and his Gestalt colleagues.

Wolfgang Köhler, the author of the most extensive early studies of "the mentality of apes" (Köhler 1925/1951), first formulated a dynamic systems approach to brain and cognition (Köhler 1920). A good deal of what S&K describe as "new" observations of coregulation, dyadic interactions, creativity and mutual adjustments, and of the social background of action and interaction, are modern replica of what Köhler described in his famous studies.

Gestalt psychologists opposed the idea of internalization of the "outer world" and worked out an original and extremely fertile concept of unity of mind and the physical world. This concept was most explicitly formulated in Köhler's theory of *physical* and *phenomenal* forms which applied, in the same terms, to the realm of mind and to the "outer world." The Gestaltists stressed the physiognomic character of perceptual experience, where physiognomic refers both to the *expressive* character of percepts (as though dynamically expressing "an inner form of life"; Werner 1957, p. 69) and to the conative dimension of perception whereby the *readiness for action* imbedded in perceptual experience "urges" us to act upon, or use, perceived objects. Perception, action, and expression were thus assumed to be closely intertwined, making up a unified "practical field" (see Rosenthal & Visetti 1999; 2003). It is precisely this unitary view of perception, action, and expression that may be requisite for semiotic interaction, forming a practical basis for communication since whatever apes, or we, perceive and do is intrinsically expressive. Yet, in order to communicate, it is necessary to have something to transmit. Interactional synchrony illustrated by the dance metaphor falls short of this requirement because it may amount to simple harmonic adjustments between dynamical systems. The need to communicate may arise instead when synchrony breaks down – that is, when dissynchrony occurs and generates instability. Moreover, we do not merely communicate in order to get hold of, point to, or share a novel experience or object, but also with the aim of stimulating the ongoing exchange or simply adding an episode to the history of our selves.

What does it mean to have language, then? Although Gestalt psychologists failed to properly appreciate the importance of language and categorization in the making of human mind, Köhler perceptively noted that it was the absence of language that prevented his most clever apes from *giving an elaborate structure to time*. Language provides the only direct means for entertaining an explicit relationship to the past or remote future. In the absence of language, it would indeed be difficult to have a history, to make up a narrative, or to carry out truly elaborate programs. The emergence of language endowed the human mind with extremely rich temporality and thereby permitted, inter alia, the incomparable development of *memory*.

*Narrative*, which is inconceivable without language, has been argued to play an essential role in the making of the human mind. As Bruner put it:

Narrative deals with the stuff of human action and human intentionality. It mediates between the canonical world of culture and the more idiosyncratic world of beliefs, desires and hopes. It renders the exceptional comprehensible and keeps the uncanny at bay. . . . It reiterates the norms of the society . . . can even teach, conserve memory, or alter the past. (Bruner 1990, p. 52)

According to Bruner, narrative requires four crucial constituents to be carried out: (a) agentivity; (b) establishment and maintenance of sequential order (event structure); (c) sensitivity to the canonical and the exceptional; (d) adoption of a narrator's perspective (because narration cannot be voiceless). There is no doubt that some apes, especially Kanzi, can handle direct agentivity, and may be able to establish short sequences necessary for the execution of simple plans. But what is the extent of their capacity to express and characterize the ordinary and the exceptional? Can they explicitly adopt another being's perspective? How much of role reversal can they handle; can they cast a third party to perform an action that is not directly observable?

Although we agree with S&K that a dyad may give substance to interactional synchrony, we emphasize that a dyad does not make a society and that language use presupposes a complex *triadic* relationship. The issue is not how many individuals are involved in interaction, but the very necessity of a *third party viewpoint*, whether the third party is *present* or *absent*, *real* or *fictitious*. What gives human thought the very characteristic architecture of textuality criterially requires a third party. Grammatical devices (person, time, space, modality) universally reflect