Bodily mimesisas “the missing link” in human cognitive evolution

Zlatev, Jordan; Persson, Tomas; Gärdenfors, Peter

Published in:
LUCS: Lund University Cognitive Studies

2005

Citation for published version (APA):
Zlatev, J., Persson, T., & Gärdenfors, P. (2005). Bodily mimesisas “the missing link” in human cognitive evolution. *LUCS: Lund University Cognitive Studies*, *121*. http://www.lucs.lu.se/LUCS/121/LUCS.121.pdf

Total number of authors:
3

Jordan Zlatev, Tomas Persson and Peter Gärdenfors, Lund University Cognitive Studies, 121, (2005).

Bodily mimesis as "the missing link" in human cognitive evolution

Jordan Zlatev[1], Tomas Persson[2] and Peter Gärdenfors[2]

Jordan.Zlatev@ling.lu.se, Tomas.Persson@lucs.lu.se, Peter.Gardenfors@lucs.lu.se[1]

[1]Department of Linguistics, Lund University, [2]Cognitive Science, Department of Philosophy, Lund University

1. Introduction

It is fairly uncontroversial that there is a large gap between the communicative and cognitive systems of non-human animals and those of human beings. There is much less consensus, however, on *what* the nature of this gap is, and even less on *how* it was bridged in evolution. A common view is that it is *language* that constitutes the quintessential human feature, and if we could only explain the origins of language, we would also have the answer to the nature and source of human uniqueness (Christiansen and Kirby 2003). Our view is in some respects contrary to this. We agree that human language is qualitatively different from animal communication, and in Section 2 we outline what we find to be the essential differences between the two. However, precisely because the gap between animal communication and language is so large, it is extremely difficult to see how one could "evolve" into the other. Therefore, despite much current enthusiasm and a variety of conflicting theories, we can at present only offer more or less compelling "narratives" (Landau 1991) or, somewhat more scientifically, "constraints" (Johansson 2005) for a theory of language origins.

In this article, we hope to add an important piece to the puzzle by arguing that there is a semiotic capacity that is more basic than language that distinguishes us from other animals: *bodily mimesis*. In Section 3 we define the concept, following the lead of Donald (1991), but elaborating it in a way that makes it more precise and less susceptible to criticism. Furthermore, our definition permits the formulation of a *mimesis hierarchy* consisting of four evolutionary stages. While the first two stages are shown to be within the capacities of non-human apes (henceforth, simply *apes*), and the fourth merges with language, the third stage, or what we call *triadic mimesis* constitutes a "missing link" that can help us explain the emergence of human uniqueness as well as the origins of language.

In the main part of the article – Sections 4 to 6 – we proceed to apply the mimesis hierarchy to three cognitive-communicative domains closely connected to bodily mimesis: *imitation*, *intersubjectivity ("theory of mind")* and *gesture*. Our main method will be to analyze evidence from primatology (and to a lesser degree from child development) and attempt to identify levels within these domains that correlate with or are even constituted by the levels of the mimesis hierarchy. By investigating three socio-cognitive domains in parallel – imitation, intersubjectivity and gesture – our intention is to suggest evidence for the most likely *source* of human uniqueness. To preempt our conclusions in Section 7, our analysis suggests that capacities such as the understanding of others' intentions (Tomasello 1999) or even the sharing of intentions (Tomasello et al. in press) are not the crucial characteristic that differentiates human beings and apes. Rather we will suggest that it is the combination of spontaneous detailed imitation of actions leading to the formation of *mimetic schemas*, and the understanding of the semiotic potential of gestures, or what we will call the *communicative sign function*. We will suggest that the

Lund University Cognitive Science-LUCS 121 ISSN 1101 – 8453

co-evolutionary cycle between imitation and gesture (as well as intersubjectivity) paved the way for the evolution of language in an evolutionary process of human "self-enculturation" that parallels the Vygotskian notion of the internalization of social, interpsychological phenomena in ontogeny.

2. From animal communication to human language

In comparing animal communication and human language it is not difficult to find striking differences on a number of cognitive dimensions: First, the *signals* that constitute animal communication are mostly predetermined by the genetic makeup of the organism. Even though it is clear that genes do not directly determine behavior but set constraints on ontogenetic development, and in this sense all development is *epigenetic* (Badcock 2000), it is still possible (and necessary) to distinguish the opposite poles of the endogeneous/exogenous continuum. In this respect, animal signals, be they the quacking of frogs or the play-face of chimpanzees, can be regarded as by and large *innate* rather than *learned.* To the extent that animal signals do involve learning – for example in fixing the group-specific "dialect" of songbirds (Marler 1991) or determining the appropriate environmental trigger for the production of an innate alarm call in vervet monkeys (Cheney and Seyfarth 1990) – this learning involves cognitively simpler (and better understood) processes such as triggering, calibration and inhibition (Hauser 1996). On the other hand, neither the expression nor the meaning of linguistic *signs* is in any way genetically determined, and the same may equally well apply to grammar (Tomasello 2003), though this is somewhat more controversial. Nevertheless, as a whole, it is clear that the learning of language is based on cultural processes of considerable cognitive complexity (Bloom 2000; Sinha 2003; Zlatev 2003), and arguably on reflective consciousness (Robinson 1996; Gärdenfors 2003; Mandler 2004; Zlatev 2005b). Thus, despite the need for some qualifications, in general terms animal signals are more or less innate, while language is learned, and culturally transmitted across generations.

Another important difference compared to language is that animal signals are tied to particular circumstances and are produced more or less automatically, i.e. with very little volitional control (Deacon 1997; Hauser 1996). In contrast, human utterances are largely independent from the physical context in which they are produced, and usually refer to states and events that are not currently present, thereby showing what Hockett (1960) called *displacement,* which also allows the speaker to *inform* the hearer of something he does not know. Furthermore, while not every step in the construction of an utterance is under conscious control, the overall activity of human discourse would be impossible unless it was governed by a kind of "intermediate-term memory" that requires reflective consciousness (Donald 2001).

Looking at communication from the side of the interpreter, the response to an animal signal is a behavioral pattern that is also relatively fixed, e.g. climbing up the tree in reaction to a specific type of alarm call. In contrast, discourse interpretation is both individually flexible and collectively negotiable (Vygotsky 1978; Sinha 2003).

Moreover, even though animal signals can and do transmit information about the external environment to the interpreter of the signal (as in the case of alarm calls) and such signals have been called "referential" and even "symbolic" (Marler 1985), this is not the kind of referential or representational relationship between expression and content that is involved in language. In connection with context independence and volitional control, human reference can be *intentionally* deceptive ("Wolf!"), while the calls of the vervet monkeys can at most be functionally so (Whiten and Byrne 1988; Hauser 1996). One way to summarize the difference is to say that human language is, in general, *triadic* – it is used by one individual for explicitly representing an existing or imaginary state of affairs for another individual. On the other hand, animal signals are generally speaking *dyadic*: the animal responds to a stimulus or *cue* in the environment, causing the emission of a particular signal directed to a conspecific, without linking self, referent and addressee into a referential triangle (Sinha 2003).[1]

---

[1] The question of the referential status of animal signals may be debated, but it is significant that one of those who most contributed to their rich interpretation as "symbols" is currently seeing them in a more modest light: "More recently, Marler's experimental approach and writings on the topic of signal meaning have moved away from high-level

Finally, the units of language stand in complex internal semantic relations involving e.g. complementarity (Deacon 1997) and grammatical relations involving hierarchical structure and recursivity (Hauser, Chomsky and Fitch 2002; Givón 2002). Nothing of the kind has been shown to exist in natural animal communication, though aspects of both semantic relations and grammatical patterns are not beyond the grasp of language-taught apes (Savage-Rumbaugh and Lewin 1994).

We may conclude that the six general differences between language and animal communication discussed in this section and summarized in Table 1 define two very different types of semiotic systems. Even though one may argue that for each of the individual features there may be some forms of animal communication that would appear to be "intermediary", e.g. the flexible interpretations of chimpanzee alarm calls by Diana monkeys (Zuberbühler 2000), *no animal in the wild has anything approaching the socially transmitted, voluntarily controlled, contextually flexible, triadic semiotic system that is language.*

As pointed out in the introduction, the difficult question is to account why and how such a distinctively "odd" (from a biological perspective) communication system emerged in the first place. Proposals in the currently exploding literature on language origins vary widely on what they find to be the crucial feature of human cognition that could "bridge the gap" to language: processing of recursivity (Hauser, Chomsky and Fitch 2002; Bickerton 2003), enhancements in social cognition (Tomasello 1999; Tomasello et al. in press), mechanisms supporting the learning and use of symbolic relations (Deacon 1997, 2003). In the present quest for the Holy Grail of explaining human uniqueness that tends to focus on the origin of language different researchers tend to offer their favourite candidate for "the missing link" and do their best to support their hypothesis, often at the price of disregarding a good deal of evidence (see Johansson 2005). Our overall project similarly aims to resolve the puzzle of what makes human beings cognitively and communicatively special, but differs from the majority in (a) not focusing on language but on a skill that we argue is a prerequisite for language and (b) drawing from a large number of empirical and theoretical studies.

Table 1. Six critical differences between animal communication and human language

	Animal communication	*Human language*
1. Degree of learning	None ("innate") or highly limited	Extensive
2. Conscious control	None or highly limited	High
3. Contextuality	Tied to a particular context (stimulus setting)	Flexible, relatively independent from specific context
4. Interpretation	(Relatively) fixed response	Flexible, "negotiable"
5. Communicative relations	Dyadic: - Cue (environment, subject) - Signal (subject, recipient)	Triadic: Speaker-addressee-referent
6. Systematicity (internal relations and combinatorics)	None, or very limited	High

Animals and human beings differ not only in their means of communication, but in (apparently) non-communicative aspects of cognition such as planning, symbolic thought, "theory of mind", etc. In other words, there is a parallel between the "communicative gap" and the "cognitive gap". What is the source of this parallelism? When it comes to ontogeny, a possible explanation for the source of the "higher functions" of human beings is the classical proposal laid out by Vygotsky (1978) that the latter can be seen as resulting from *the internalization of interpersonal relations and representations*, or as expressed by Vygotsky in an often quoted passage:

> Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first between people (interpsychological), and then inside the child (intrapsychological). This applies equally to voluntary attention, to logical memory, and to the formation of concepts. All the higher functions originate as actual relations between human individuals. (Vygotsky 1978: 57)

We wish to suggest that an evolutionary analog to this process, i.e. a process of hominid self-

---

cognitive interpretations and toward more low-level mechanisms linked to the particular stimulus features of the environment." (Hauser 1996: 59) enculturation in which more complex forms of communication created selection pressures for more complex forms of cognition, can help us understand the close connections between the two "gaps". Such a socio-cultural perspective on human cognition and its evolutionary origins does not conflict with, but rather presupposes the crucial underlying role of our general primate cognition (Donald 1991; Tomasello 1999; Zlatev 2003). In this article we will adopt such a bio-cultural perspective as a general "working hypothesis", without thereby arguing against more individual-based alternatives (e.g. Gärdenfors 2003). Ultimately, we will need a synthesis of the social and the individual aspects of human cognition in order to make sense of the nature of our major theoretical concept that is both inter- and intra-personal, as we show below.

3. Bodily mimesis

The concept of *mimesis* goes back at least to Aristoteles, but as far as human cognitive evolution is concerned, we take our departure from the *mimesis hypothesis* of human origins presented by Donald (1991, 1998, 2001). This hypothesis states that a specific form of communication and cognition (and corresponding culture) mediated between those of the common ape-ancestor and modern humans based on "the ability to produce conscious, self-initiated, representational acts that are intentional but not linguistic." (Donald 1991: 168) In brief, Donald proposes that while ape culture is based on "episodic cognition" and associational learning, early *Homo* – most likely *Homo erectus/ergaster* considering the relative jump in brain size and material culture witnessed at this stage of hominid evolution – evolved a new form of cognition. What supported this was the fact that the body could be used volitionally to do what somebody else is doing (imitation), to represent external events for the purpose of communication or thought (mime, gesture) and to rehearse a given skill by matching performance to a goal.

The most important features that Donald attributes to mimesis are: (a) *reference*: mimetic representations stand for something (for someone); (b) *intentionality*: this referential relationship is grasped both by the "mimer" and the interpreter; (c) *communicativity*: the relationship in (b) can be used for the purpose of conveying thoughts; (d) *autocuing*: production is voluntary as opposed to instinctive; and (e) *generativity*: the "ability to "parse" one's own motor actions into components and then recombine these components in various ways" (Donald 1991: 171).

It can be observed that features (a)-(e) define a system of communication that is intermediate between animal communication and language, as characterized in Section 2 and Table 1. Like the latter it is learned, flexible and possibly triadic. At the same time, it lacks the following critical features of human language:

- *Full conventionality*: not just shared, but known to be shared and thus a part of mutual knowledge (Itkonen 1978; Clark 1996);
- *Arbitrariness*: the semiotic ground for the expression-content relation requires neither similarity, nor contiguity (Peirce 1931-1935; Sonesson in press)
- Extensive *systematicity* of the internal relations among signs (Deacon 1997; Zlatev 2003).

Thus the hypothesis that mimesis plays the role of a "missing link" (in the words of Donald himself) in human cognitive evolution is *a priori* attractive. Furthermore, it has been backed up by evidence from archeology, neurobiology, cognitive psychology and developmental psychology, such as the homology between "mirror neuron" systems in monkeys used for action recognition, and structures for the control of imitation, mentalizing and language in human beings, the presence of a "mimetic stage" in human ontogeny and the presence of mimetic thought in patients with total aphasia (e.g. Donald 1991, 2001; Nelson 1996; Zlatev 2002, 2005; Corballis 2002).

Nevertheless, the mimesis hypothesis has been met with much resistance and criticism (Mitchell and Miles 1993; Tomasello 1993; etc) that can be summarized as follows: On the one hand, Donald's theory underestimates the cognition of non-human primates with respect to tool-making and cultural traditions (Boesch 2003) and their ability to understand others' intentions (Tomasello, Call and Hare 2003). On the other hand, features (a)-(e) attribute so much representational complexity to mimesis, that this obviates the need for a second cognitive transition to language (Laakso 1993). If this criticism is correct, then mimesis can hardly "bridge the gap" as required, since it, so to speak, both gives too little to the apes and too much to *Homo erectus*, making the width of the gap too large to bridge. Finally, Donald's theory does not address the question why mimesis evolved in the first place, and why it later became superseded by language.

Appart from these issues, Donald's definition of mimesis is unclear with respect to at least the following points: Are the features (a)-(e) mentioned above necessary and jointly sufficient or are some more central than others? Is there a priority for the communicative or the non-communicative forms of mimesis? In which way should one interpret the notoriously ambiguous terms "reference" and "intentionality"? Furthermore, if mimesis possesses all the features of language except for phonetic realization as claimed by Laakso (1993), it remains unclear what exactly makes it "not linguistic" in the words of Donald (see the citation given at the beginning of this section). Hence, our first step in elaborating the mimetic hypothesis of human origins is to offer a more specific (re)definition of the concept, which we refer to as *bodily mimesis* in order to distinguish it from the broader concept with Aristotelian roots:

(Def) Bodily mimesis: A particular act of cognition or communication is an act of bodily mimesis if and only if:

1. It involves a cross-modal mapping between proprioception (kinaesthetic experience) and exteroception (normally dominated by vision), unless proprioception is compromised. (**Cross-modality**)
2. It it realized by bodily motion that is, or can be, under conscious control. (**Volition**)
3. The body (part) and its motion correspond to – either iconically or indexically – to some action, object or event, but at the same time are differentiated from it by the subject. (**Representation**)
4. The subject intends for the act to stand for some action, object or event for an addressee. (**Communicative sign function**)

But it is not an act of bodily mimesis if:

5. The act is fully conventional, i.e. a part of mutual knowledge, and breaks up (semi)compositionally into meaningful subacts that relate systematically to each other and to other similar acts. (**Symbolicity**)

This definition resolves some of the problems mentioned earlier. By stating that bodily mimesis involves a mapping between *proprioception*, involving kinesthetic sense and the input from skin receptors and serving as our basic form of self-knowledge (e.g. Edelman 1992; Gallagher 1995, 2005) and exteroperception (above all vision, but also hearing and touch), this explains the sense in which bodily mimesis involves the body, including specific organs such as the vocal tract – to the extent that they are involved in proprioceptive-exteroceptive mapping. Since speech production, and possibly even speech comprehension according to the "motor theory of speech perception" (Liberman and Mattingly 1985) involve such a mapping, this implies that even speech involves bodily mimesis, at the same time as it transcends it due to symbolicity. By focusing on full conventionality and systematicity as distinguishing language from bodily mimesis, it becomes clear in which sense mimesis is "not linguistic". Most importantly, however, the five criteria in the definition of bodily mimesis presented above allow us to classify human and ape socio-cognitive skills along a scale consisting of at least four distinct stages. We refer to this as the *mimesis hierarchy* (see Table 2).

Table 2. *The mimesis hierarchy*. Definitions of the four evolutionary stages (for terms in italics, see the definition of bodily mimesis in the text) and examples of corresponding types of acts.

Stage	*Definition*	*Examples*
1. Proto-mimesis	A bodily act involving *Cross-modality* with proprioception, but lacking *Volition* or *Representation* (or both)	Facial expressions, bodily synchronization
2. Dyadic mimesis	An interpersonal or intrapersonal bodily act displaying *Volition* and *Representation*, but not *Communicative sign function*	Shared attention, imperative pointing, mirror self-recognition, do-as-I-do imitation
3. Triadic mimesis	As (2) but also involving *Communicative sign function*	Joint attention, declarative pointing, pantomime
4. Post-mimesis	As (3), but also involving *Symbolicity*	Signed language

In the remainder of this article we shall apply the proposed mimesis hierarchy to the evolution (and to a lesser degree, the ontogeny) of three cognitive-communicative capabilities: *imitation* (Section 4), *intersubjectivity* (Section 5), and *gesture*s (Section 6). By analyzing evidence from primatology, and to a smaller extent child development, we will identify levels within these domains that correspond to the levels of the mimesis hierarchy. In other words, we consider whether there are simple forms of these capacities that are in essence proto-mimetic; whether there are forms that are dyadic mimetic; whether there is furthermore evidence for communicative intentions, and thus for triadic mimesis. Finally we ask whether there are specific forms of imitation, intersubjectivity and gesture that depend on the acquisition of a conventional symbolic system, i.e. language, and are thus post-mimetic.

We will show that the progression *proto-mimetic* > *dyadic-mimetic* > *triadic-mimetic* > *post-mimetic,* instead of the dichotomy animal signaling vs. human language outlined in Section 2, and in particular by distinguishing between dyadic and triadic mimesis, has implications for a theory of the origins of human cognitive uniqueness. It may seem that such a theory goesagainst the stream of much current theorizing in evolutionary science since it invokes relatively discrete "stages" requiring qualitative transitions, while both the nature of evolutionary processes involving many small mutations and the available fossile evidence seem to speak in favor of a prolonged gradual process (Johansson 2005). However, we hasten to note that there is no contradiction between the stage-like theory that we will pursue and the evolutionary evidence. Firstly, it is completely possible to have relatively discrete transitions, or "punctuated equilibria" (Eldredge and Gould 1972) within a basically gradual framework, i.e. without this implying any form of "saltations" or sudden "macro-mutations" (Dessalles 2004). The likelihood for *fairly* rapid and abrupt changes is even higher when biological and cultural evolution interact, as has most likely been the case with human cognitive evolution including language (Deacon 1997). Secondly, our notion of a mimesis hierarchy, as other similar constructs (Dennett 1996; Zlatev 2003), is meant to serve primarily as a *conceptual role* in helping us relate and compare primate skills in somewhat different domains: it is not an evolutionary ladder for classifying different species, which may very well have different degrees of the mimetic skills under discussion, and thus form much more of a cline. With these caveats we procede with our analysis.

4. Mimesis and imitation

Chimpanzees, bonobos, gorillas and orangutans can copy behavior to some extent and in seemingly various ways. The literature on ape imitation in the broad sense of the term (see Whiten et al. 2004 for a recent review) can for our purposes be grouped into the following main experimental paradigms:

- Neonatal mimicking
- Instrumental tool-use
- Experiments with "artificial fruits" (puzzle boxes)
- "Do-as-I-do" tasks
  - in respect to body movements ("Simon says")
  - in manipulation of objects
- Ethological methods (i.e. field observations of food processing techniques or indirect evidence from local traditions)

Most obviously relevant for our discussion of stages of bodily mimesis are those types of observational learning in which *an observer copies something of a model's bodily actions*, may it be in a laboratory tool-using situation, natural foraging situation or mimicking in a do-as-I-do game. Other forms of observational or social learning, like *emulation*, available in several versions (Byrne 1998; Custance, Whiten and Fredman 1999), *response facilitation* (see Byrne 1999), and *stimulus- and local enhancement* (see Tomasello 1996), can inform observers about properties of the environment and objects, or the connection between certain manipulations and outcomes, but do not involve copying bodily actions *per se* and thus it is not clear if they involve bodily mimesis.

For our purposes, the collective term *imitation* pertains most clearly to the copying of bodily motions (the do-as-I-do experimental paradigm) as well as to imitation on the *action level* in object manipulation. The notion of action level imitation was originally contrasted to that of *program level* (Byrne and Russon 1998), the latter being the hierarchical organization of a series of actions and the former the execution of an individual component action. Note that imitation on the program level can take place without imitation on action level. The program level has been further divided by Whiten (1998) into being either of a sequential or a hierarchical kind. An example of *sequence imitation* from the artificial-fruit experiments would be to manipulate obstacle A before obstacle B, while an example of imitation on the action level would be poking vs. pulling to remove a plug. If the subject copies the manual movement of the model, it is regarded as an instance of action-level imitation, but if the ape instead copies the movement of the plug, it is categorized by Custance et al. (1999) as *object movement re-enactment*. The latter is usually regarded as a form of emulation, where the end state of an event is achieved (or copied) but the subject invents his or her own way to achieve this end and the model is in principle superfluous. However, since object movement re-enactment is a special kind of emulation where at least the direction of the movement of the object is copied, in our analysis it would qualify as a form of bodily mimesis.

There is also the hypothetical case where the subject only learns that the plug can be moved and then proceeds to bring about that effect in whatever way he or she prefers. To determine what is what, however, is a tricky business. In the latter case the subject might happen to bring about by chance the same movement of the plug as the model showed, and thus give the impression of object movement re-enactment. Furthermore, in both a false and in a true case of object movement re-enactment, the subject can also happen to perform the same manual actions as the model by chance and give the impression of bodily imitation on the action level. At least three forms of observational learning can thus yield the same expression: simple emulation, object movement re-enactment and action level imitation. The experiments with artificial fruits, some of which we will review below, are designed to tease apart these copying strategies.

The failures of great apes to fully copy behavior in several early studies led Tomasello, Kruger and Ratner (1993) to conclude that apes emulate, but cannot imitate, since imitation in their view entails copying both of the goal and the method to bring about that goal. For Tomasello and his collaborators imitation thus has a "theory of mind" component by definition, and on the basis of the imitation data, the group has argued that apes lack an understanding of others as intentional agents (e.g. Tomasello 1999). Since then both the field of primate imitation studies has expanded and the MPI Leipzig group has modified their views on apes' abilities to understand intentional agency (Tomasello et al. in press, see Section 5 for more discussion).

Donald's definition of mimesis (presented in Section 3), comprising *reference*, *intentionality*, *communicativity, autocuing* and *generativity*, is difficult to apply to non-human primate imitation because observational learning through copying has been studied in non-communicative tool using and puzzle box situations, with few exceptions such as the study by Tomasello et al. (1994) on chimpanzee gestures in which the authors did not find any observational learning involved in functional gesture acquisition. However, using our definition (see Section 3), it is possible to apply the concept of bodily mimesis to findings in the field of ape imitation. As we will show in this section, these findings can be interpreted as evidence that apes are capable of at least proto-mimetic and dyadic mimetic imitation, and possibly even triadic mimetic imitation when a *communicative sign function* is added.

4.1 Proto-mimetic imitation

In what we call proto-mimetic imitation there is (iconic) resemblance between the copy and the copied, which is probably based on a basic form of identification with the model (see Section 5.1). For this to be classified as proto-mimetic rather than dyadic-mimetic, there should be no clear evidence for volition or for differentiation between oneself and the model.

The prime candidate for proto-mimetic imitation is neonatal facial mirroring. In the case of human neonatal mirroring, the field is divided between those who prefer a minimalist interpretation and those who argue for a more mentalistic one. For example, Anisfeld (1996) in a meta-analysis of experiments by Meltzoff and Moore and Heimann argues that the infant's response is limited to one facial movement, tongue protrusion, and not several. If so, it would be a largely reflexive form of mirroring unlikely to require *volition* on the part of the infant, and thus would be proto-mimetic. On the other hand Meltzoff and Moore (1977, 1983, 1994) have argued that the human infant not only displays volition (i.e. choice on whether to imitate or not, and ability to mirror more than tongue protrusion), but that he/she *differentiates* between him/herself and the model. Meltzoff and Moore (1994) argue that there is evidence for deferred imitation in 6-week-old infants who imitated the action when seeing the model again after a delay of 24 hours. While differentiation in itself does not require the kind of displacement that is witnessed in deferred imitation (see Sonesson in press), we agree with Piaget (1945) that the capacity for deferred imitation rests on at least a simple form of representational activity. Thus, if the Meltzoff-Moore interpretation of neonatal imitation is correct, we must conclude that human children perform dyadic imitation (see below) from a very early age.

In the case of apes, newborn chimpanzees have been shown to have a period of innate mirroring responsiveness to human facial stimuli (mouth opening and tongue protrusion) that is present to 11 (Myowa-Yamakoshi 2001) or 9 (Myowa-Yamakoshi et al. 2004) weeks of age. This shows that neonate chimpanzees have (at least) the capacity for proto-mimesis with respect to faces. It is to our knowledge not yet tested whether this capacity also involves differentiation or volition. But different sources for highly similar responses in human and ape newborns seem unlikely.

4.2 Dyadic-mimetic imitation

In what we call dyadic-mimetic imitation, the subject differentiates between oneself and the model and understands that a bodily posture or motion can correspond to something else, like an object or action. However, there is no attempt to communicate this "something else" to another individual by means of the bodily motion.

In a non-communicative setting, like in the experiments on social learning studied with non-human primates the "something else" mentioned above would be to understand the meaning of a model's actions: what the model is trying to achieve. As mentioned above, imitating the goal as well as the means to bring about that goal is the hallmark of (true) imitation according to Tomasello, Kruger and Ratner (1993), and involves certain levels of intersubjectivity involving understanding of intentions.

When trying to implement the imitation of goals in an experimental setting the goal is often to get hold of a food item, and these experiments have tended to either foster innovation or use of old reliable methods. This is not surprising. If you truly want to get hold of a food item, and apes clearly do so, the most effective means one can come up with is used, and that is not necessarily the one witnessed.

The well-known rake experiment of Nagell et al. (1993) is an example of imitation studies using an instrumental tool. If there is any learning involved in the case of apes it is generally considered to be in the form of emulation: the subjects learn something new about tools, their movements or the environment, but not about the bodily actions the model is using to bring about the intended goal.

Better imitation results have been obtained in experiments with so called "artificial fruits", i.e. boxes with obstacles that the subject has to work around in order to get to the food content. These experiments address the question: What do apes spontaneously copy when they get an opportunity to observe a model? Which strategies and mechanisms of social learning are involved? The artificial fruit experimental paradigm has yielded mixed results, as strategies of result emulation, object movement re-enactment, action-level imitation and sequence imitation have been recorded (Whiten et al. 2004). The results seem to vary over fruit used and species tested and/or age of subjects (Stoinski and Whiten 2003; Custance et al. 2001; Stoinski et al. 2001; Whiten et al. 1996). Action-level imitation implies that the subject can represent a model's manual movements and re-enact these from memory. Performing *sequence imitation* means that the subjects are also able to parse a string of component actions and later reassemble these in the same fashion from memory, which again show their ability to re-enact detached representations voluntarily.

Apes have demonstrated some skills of action-level and sequence imitation, but at the same time they experience some difficulties with them as evident by their lack of full copying. Imitating the use of familiar motor actions in novel situations seem to come about more easily in apes than copying new motor actions altogether (Myowa-Yamakoshi and Matsuzawa 1999). Especially for young individuals actions are more likely to be copied if they are close to the ones already in the subjects' repertoire (Bjorklund et al. 2002). It is not yet settled whether apes' relatively poor performance on detailed imitation of bodily actions is due to motor-perceptual aspects or skills in intersubjectivity. However, since recent studies have shown that at least chimpanzees have a degree of understanding of the mentality of others (see below and Section 5.2), the motor-perceptual explanation should not be underestimated.

Goal directedness is lacking in the most bodily mimetic paradigm, the "do-as-I-do" experiments, where part of making sure that the copy is acquired through observation is to perform never-before seen, or at least unusual, movements without any practical function. There is no food item to get to at the end of the action sequence. The subject is verbally asked to do the same as the model on cue. The do-as-I-do paradigm is an example of studies that ask the question: *Can* apes copy behavior? The answer is positive. Both cross-fostered (or "enculturated") apes, e.g. the orangutang Chantek (Miles et al. 1996) and the chimpanzee Viki (Custance et al. 1995), and chimpanzees with more moderate human upbringing (Custance et al. 1995) can pass these tests. Once they understand the meaning of the "do the same" command, any faults in the copy is bound to be motor-perceptual and not due to intersubjectivity issues. The do-as-I-do experiments prove that apes are capable of both *volition* and *representation* (they represent the model's body), which qualify them as dyadic-mimetic imitators. However, learning the concept of SAMENESS may influence the performance of apes profoundly. Thompson, Boysen and Oden (1997) found that language-naïve chimpanzees learned to judge relations between relations (A is to A as B is to B, and not as C is to D) in a matching-to-sample task only if the subjects had previously learned a token for the concept SAME. Ordinary matching-to-sample training could not yield the same result. Experiments that hinge on the subject "getting the point of the game", so to speak, have implications for ecological validity since the experimenter has to plant certain knowledge in the subject that it might not have stumbled upon naturally. This is especially problematic for evolutionary reasoning around comparative experimental psychology, since it is not clear what circumstances would have "planted" such knowledge in the hominid line.

Deferred imitation on objects, involving do-as-I-do imitation, is perhaps the most promising evidence for dyadic mimetic capacities in apes. In these experiments the subject is shown actions on objects and is then either asked to "do the same thing" (if the the verbal command is learned) or (if not) the spontaneous handling of the object is recorded. The results show mainly approximate imitation but also exact copying and a notable increase in ability with age (Bjorklund and Bering 2003; Bjorklund et al. 2002; Bjorklund, Bering and Ragan 2000; Bering et al. 2000). However, mother-reared chimpanzees seem to do less well while enculturated apes can outperform human children on certain tasks (Tomasello, Savage-Rumbaugh and Kruger 1993). The fact that the test is made after a delay heavily strengthens the need for *representation* and recall and execution from memory. Bjorklund and Bering (2003) suggest that part of what comprises the enculturation phenomenon is a greater conceptualization of human behavioral programs. It is possible that non-human primates in wild populations have critically less exposure to long series of goal-directed behavior than apes around humans to fully be able to follow through longer strings of actions.

Byrne (1999) makes a similar proposal, but stresses that it is general familiarity with the component actions in human manipulation of objects that gives enculturants the relevant advantage. A third possibility is that deferred imitation on objects has been tested without any other reward than what the object manipulation itself can give, and that enculturated animals find human objects more interesting and rewarding than do others. Infant apes seem to be more interested in objects when objects are handled by human caretakers interacting with the infant than when the objects are handled by their conspecific mothers, who typically do not engage infants with objects (Bard and Vauclair 1984).

It has furthermore been found that the direction of object manipulation, i.e. an object's movement in space, is more often copied than bodily actions performed on the object (Myowa-Yamakoshi and Matsuzawa 1999). This again implies that apes copy the most salient rather than detailed aspects of an action. It is also possible that since objects draw much attention it might obscure the relevance of a model's manual actions, although this remains an untested hypothesis to our knowledge.

Whiten et al. (2004) conclude in their review of the imitation literature that many experimental results on imitation can be interpreted as object movement re-enactment (a form of emulation) rather than copies of bodily actions. But the line between the movement of an object and the hand applying the force to the object in order to bring about that movement is not a sharp one. (Does one copy the fall of the hammer or the fall of the hand holding it?) Thus, we see no reason to treat both of these types of imitation – action level imitation and object movement re-enactment – as being qualitatively different. This is captured in our analysis by stating that both involve dyadic mimesis.

Finally, social learning of feeding techniques in gorillas (Byrne and Russon 1998) and chimpanzees (Stokes and Byrne 2001; Byrne and Stokes 2002), and human chore re-enactment (e.g. washing clothes) in orangutans (Russon and Galdikas 1993) are further behaviors that might fall in the dyadic mimetic category, although in the latter case it is difficult to argue that the re-enacted act represents the observed one, since the function of the act is not the same: An orangutan does not re-enact cleaning when mopping a floor with dirty water. At the same time, as in the "Simon says" game, the orangutan can still be said to *represent* the bodily actions themselves, or if it is a case of object movement re-enactment, the movements of objects.

4.3 Triadic-mimetic imitation

Triadic-mimetic imitation involves the full-blown form of bodily mimesis, with the characteristics *volition*, *representation* and *communicative sign function*. Triadic mimesis introduces a communicative context that is mostly lacking in imitation studies on apes since these studies seldom take into account both parts of the learning dyad: model and imitator. Teaching, in the form of instruction, seems to be absent in ape interactions (Boesch and Tomasello 1998; Matsuzawa et al. 2001).

Above all, triadic bodily mimesis strengthens the call for understanding (others') mentality, which is perhaps the most problematic issue when it comes to relating bodily mimesis and imitation, mainly because the role of intentionality in imitation is already debated. If we were to follow Tomasello (Tomasello, Kruger and Ranter 1993; Tomasello 1996, 1999) in assuming that apes lack second-order intentionality ("I realize that you have intentions"), and that they can therefore not imitate but only emulate, apes should necessarily lack mimetic abilities of a triadic kind. However, the most recent and promising experiments on apes' perspective-taking have shown that chimpanzees know the importance of others' visual attention, both in communicative contexts (Tomasello et al. 1994; Krause and Fouts 1997; Pika, Liebal and Tomasello 2003, 2005) e.g. use of attention-getting gestures and competitive contexts (Hare et al. 2000; Hare, Call and Tomasello 2001). Apes also seem to be able to distinguish intentional from accidental acts (Call and Tomasello 1998) and to judge an experimenter as either unwilling or unable to give them a treat (Call et al. 2004). The conclusion is therefore that apes *do* understand second-order mentality, at least in the form of second-order attention and (possibly) second-order intentions. However, understanding the communicative sign function, and thus triadic mimesis, requires third-order mentality – as we will argue in Section 5 – and there is yet no clear evidence that (non-enculturated) apes are capable of this.

Possible instances of triadic mimetic imitation in non-human primates would be copying of referential signs in language-taught apes. However, most of the learning of signs by apes that have been taught a signed language, such as the gorillas Koko and Michael (Bonvillan and Patterson 1993, 1999), the chimpanzee Washoe (Fouts 1972) and the orangutan Chantek (Miles 1990) takes place by getting the hands molded into the signs and therefore does not qualify as either imitation or bodily mimesis. Still, Gardner and Gardner (1969) also used the "do-as-I-do" paradigm to teach new signs to Washoe with some success. A more general problem, however, is that it is difficult to distinguish the imitative learning of such signs from post-mimesis since these apes were taught a simplified form of American Sign Language (ASL), and that implies that they were exposed to, and possibly even acquired at least a degree of symbolicity (involving conventionality and systematicity). In a study of the functions of the *repetition* of lexigrams by language-trained chimpanzees and bonbobos Greenfield and Savage-Rumbaugh (1993) report at least some clear cases of the acquisition of lexigrams through imitation. At the same time

Savage-Rumbaugh and Lewin (1994) emphasize that only when the chimpanzees Sherman and Austin acquired the referential meaning of the lexigrams, after a prolonged and laborious process of training, could they match lexigrams to samples.

Perhaps the most convincing evidence for triadic-mimetic imitation in apes involves the *invention* of iconic signs by Koko (Patterson 1980) and Chantek (Miles 1990). Such sign are examples of *representation* (the ape represents something in the environment) and *communicative sign function* (the ape addresses this to his or her caregivers). Since the imitation is creative, involving the resemblance (iconicity) between the sign and the way the referent is typically used, it cannot be solely a product of being taught a symbolic code. The fact that Koko's and Chantek's sign inventions build on iconicity suggest that they are cases of triadic-mimetic rather than post-mimetic imitation. We return to these cases in Section 6.

4.4 Post-mimetic imitation

While speech may be regarded as in part post-mimetic (see Section 3), the paradigm example of post-mimesis is signed language: it is based on cross-modal mapping between proprioception and vision, and possesses the other features of mimesis: volition, representation and communicative function. Furthermore, it is *symbolic*, not in the Peircian sense of lacking any motivational relationship between the expression and content poles of its signs – over 50% of the signs of ASL are judged to be iconic (Woll and Kyle 2004) – but in the sense of consisting of *conventional* signs (presupposing third-order mentality, see Section 5) and systematicity (i.e. internal relations between the signs). Other instances of post-mimesis would be *emblems* like the thumbs-up "OK" sign, or nodding instead of saying "yes", which fulfill the property of conventionality, and at least a limited form of systematicity.

Accordingly, post-mimetic imitation would consist of skills involved in the learning of such signs. As pointed out above, some apes that have been taught a signed language did acquire some of their signs in this manner, though the majority of signs were acquired through molding of the hands. Even for those that were acquired by imitation, there may be doubt concerning the degree of their conventionality and systematicity, e.g. deaf native signers have difficulties interpreting apes' signing (Pinker 1994).

One form of imitation that is clearly symbolic (and hence post-mimetic) is what Tomasello (1999) calls *role-reversal imitation*: imitation of sign use in which the subject learns a new sign by observing and imitating the signer, and at the same time shows understanding of the sign's meaning by "reversing" the perspective to fit his role in the interaction. A simple case would be to learn the sequence "Here you are – Thank you!" so that after a while even the child can offer an object to the adult with words "Here you are" (or equivalent) and expect to hear "Thank you". A child acquiring a signed language performs likewise, reversing e.g. the direction of signs like GIVE, COME and the referent in deictic signs like I and YOU, which involve pointing to the body of ego and the interlocutor, respectively. In one of the few well-documend cases of the spontaneous emergence of gestural communication in apes, Tanner and Byrne (1996, 1999) and Tanner (2004) show that at least one gorilla, the male Kubie, uses deictic gestures to refer to himself, without his interlocutor, Zura imitating these in a role reversal manner (see Section 6.3 for more discussion of these gestures). In the case of language-taught apes, self-referential gestures are (typically) taught by molding, rather than imitating (Patterson 1980). Hence we may conclude that apes, including those who have been taught language, do not display post-mimetic imitation in a non-ambiguous manner.

4.5 Summary and conclusions

In this section we analyzed evidence from the literature pertaining to apes' imitation in terms of the mimesis hierarchy defined in Section 3. Some of the various types of imitation discussed were classified according to our model as shown in Table 3. The conclusion is that while apes are clearly capable of proto-mimetic and even dyadic mimetic imitation, it is less clear if they are capable of triadic mimetic imitation (related to their difficulty in comprehending the communicative sign function). Finally, post-mimetic imitation is beyond their competence.

Table 3. The mimesis hierarchy and types of imitation

Level of mimesis	*Imitation skills*
Proto-mimesis	Neonatal mirroring
Dyadic mimesis	Do-as-I-do imitation Object movement re-enactment Action level imitation
Triadic mimesis	Learning sign use through imitation Invention of iconic signs (rare cases)
Post-mimesis	Role-reversal imitation (not attested)

This evidence supports some other conclusions as well. Although apes have been shown – in a number of different paradigms – to be able to copy actions and bodily movements through observation, a general trend is that they make *proximate* rather than exact matches, and sometimes copy on a hierarchical or sequential level rather than on the more basic action level where the actual bodily matching takes place. This is in stark contrast to some human children who tend to "over-mimic" on the action level on the same artificial-fruit tasks as apes (Whiten et al. 1996) or copy a way to use a tool that is ineffective for the goal (Nagell et al. 1993).

Apes also tend to focus on and re-enact the movement of objects rather than body movements. This, of course, can be argued to be the logical strategy when dealing with the mechanics of objects. Approximations in matching have implications for our *volition* and *representation* requirements since they suggest that apes do not possess skills that are specialized enough to match their movements against a template, which is a prerequisite for mimetic signing. This has sometimes been interpreted as emulation caused by lack of theory of intentions (e.g. Tomasello, Kruger and Ratner 1993), but can also, and probably more likely, be ascribed to motor-perceptual difficulties per se. Also, as pointed out earlier, there is growing evidence that apes are capable of understanding others' mental states.

In our framework, both sorts of limitations can be argued to involve bodily mimesis: (a) with respect to forming mimetic representations (schemas), and (b) even more so of communicating these representations to another. In contrast, the "generative" human ability to parse series of actions (regardless of understanding their function or the model's intentions) and to re-enact these seems to be available to apes. In other words, and relevant for the debate on the origins of language, it is not so much the *syntax* of imitation that is difficlt for apes, but its *semantics*. It is conceivable how ape copying ability, coupled with a motivation to *share* one's mental states (which can, of course, also mean attempting to impose one's view on others), can give rise to mimetic-communicatory interactions, in which levels of mimesis and intersubjectivity are closely connected, as we will argue in the following section.

5. Mimesis and intersubjectivity

We understand the notion of *intersubjectivity* broadly (and rather literally) as *the sharing and understanding of others' mentality*. The term "mentality" is taken here to involve not only beliefs and other forms of *conceptual* knowledge, but all forms of consciousness, including emotions, desires, attentional foci and intentions (see Gärdenfors 2003). Not all of these need to be understood conceptually, but may involve pre-conceptual, non-reflective understanding.

While the close connection between bodily mimesis and imitation discussed above, and gesture (to be analysed in Section 6) is fairly straightforward, this may not be the case with respect to intersubjectivity. The reason for this lies, we believe, in the dominant cognitivist view of mentality as consisting entirely (or at least predominantly) of propositional representations. In this view, to understand another's mind would imply having beliefs concerning their beliefs. This conception can easily be framed as having some sort of "theory of mind" and, indeed, this has been the dominant conception in the field (e.g. Baron-Cohen 1995).

A great deal of the debate about the cognitive abilities of apes has concerned *whether or not* they have a "theory of mind" (Woodruff and Premack 1979; Byrne 1995; Gómez 1994; Heyes 1998; Tomasello 1999). Also in the discussion of human cognitive development, the question *when* and *how* children develop one has been a central focus of content (Perner et al. 1987; Gopnik and Astington 1988; Gopnik and Meltzoff 1997; Mitchell 1997).

But regarding intersubjectivity as a "theory" involving (primarily) beliefs has had some negative implications for understanding this fundamental suite of socio-cognitive capacities: (a) the question is typically posed as involving a discrete, *all-or-nothing property*: either the child or animal "has a theory of mind" or doesn't; (b) the role of one's own *body* and *actions* – and those of the other – is minimized, and (c) "theorizing" about the minds of others (or even oneself) is considered prior to *acting* in consort with them, giving intersubjectivity a rather static, detached quality.

Recently, however, these implications have been seriously questioned on the basis of both new empirical evidence from child development and primatology and older philosophical, above all phenomenological, insights. With respect to (a) it has become clear that "the generic term 'theory of mind' actually covers a wide range of processes of social cognition" (Tomasello, Call and Hare 2003: 239). Consequently one must distinguish between these different processes in order to understand the cognitive capabilities of animals, as well as children of different ages. For example, Gärdenfors (2003) splits the so-called "theory of mind" into six levels, where the middle four clearly correspond to different socio-cognitive capabilities: understanding others' emotions, attention, intentions and beliefs, respectively. Such a division can be complemented with the hypothesis that the different capabilities correspond to different phylogenetic and ontogenetic *stages* in evolution and development, and indeed Gärdenfors (2003) makes this proposal.

The criticism concerning (b) and (c) was developed first within the phenomenology of Husserl some one hundred years ago and targeted the intellectualist perspective on the "understanding of other minds" (the usual term for intersubjectivity in philosophy) as a matter of inference or analogy from the knowledge of one's own mind. As summarized by a modern interpreter: "For Husserl, understanding another person is not a matter of intellectual inference, but a matter of sensory activations that are unified in or by the animate organism or lived body that is perceiving another animate organism." (Gallagher in press).

This perspective on intersubjectivity was further developed by Merleau-Ponty (1962), in particular through the notion of the "corporeal schema" which serves as "a normal means of knowing other bodies" (Merleau-Ponty 2003: 218). More recently, Gallagher (1995, 2005, in press) has elaborated the distinction between *body schema* and *body image* where the first is preconscious and serves as a precondition and backdrop for intentionality (cf. Searle's (1992) notion of the Background), while the latter is "a (sometimes conscious) system of perceptions, attitudes, beliefs and dispositions pertaining to one's own body" (Gallagher 2005: 37). Empirical neuroscience has recently given support for the role of the body in understanding others through mechanisms involving "mirror neurons" (Rizzolatti et al. 1996; Gallese, Keysers and Rizzolati 2004; Arbib in press), that mediate between the perception of another and the subject's own proprioception and action.

Within this "embodied" and action-oriented perspective on the understanding of other minds, unlike in the intellectualist cognitivist one, bodily mimesis becomes clearly relevant. With respect to distinguishing different levels of intersubjectivity, we can apply the mimesis hierarchy defined in Section 3 to yield a possible evolutionary and developmental hierarchy of intersubjectivitity.

It should be noted that this way of "splitting the theory of mind" is different from that proposed by Gärdenfors (2003) since the different capabilities distinguished by the Gärdenfors (understanding emotion, attention, intention and belief) are in a sense orthogonal to the mimesis hierarchy. For example, the understanding of others' emotions occurs on each one of the levels of the mimesis hierarchy, with increasing cognitive complexity. However, if we consider what may be regarded as a *typical manifestation* of the respective capability (emotion, attention and intention), there turns out to be considerable overlap between the levels of intersubjectivity defined in this article and those of Gärdenfors (2003): Emotion can be shared on a proto-mimetic level, sharing attention involves (at least) dyadic mimesis, understanding *communicative* intentions is centrally implicated in triadic mimesis (but "simple" intentions are understood dyadically), and understanding beliefs appears to require explicit symbolic skills, which can be argued to be post-mimetic since they depend on language.

A central question in relating the proposed mimesis hierarchy and intersubjectivity is whether the respective mimesis level serves as a *precondition* and a *causal factor* for the development of corresponding skills of intersubjectivity. Or is it rather that independently reached insights into the mind of others makes increasingly complex forms of mimesis possible? Our general Vygotskyan approach suggests the first scenario: bodily mimesis is a fundamentally interpersonal activity that excercises (in ontogeny) and provides selection pressures (in phylogeny) for developing more refined skills in mind reading. Thus, our predominant take is that mimesis drives intersubjectivity rather than the other way around. At the same time we acknowledge that the "driving" metaphor is not completely adequate since the causality runs in both ways and is possibly best described in terms of *co-evolution*. We will return to this issue in the summary at the end of this section.

5.1 Proto-mimetic intersubjectivity

Given our definition of bodily mimesis we can regard some of the most basic forms of intersubjectivity as "proto-mimetic" to the extent that they (a) consist of forms of interpersonal interaction that involve cross-modal mapping between proprioception and the (visual) perception of others, (b) do not fulfill either one of the characteristics *volition* and *representation*. This can be made more precise using the distinction made by Gallagher (1995, 2005) between *body schema* and *body image*. The first is characterized as "a system of sensory-motor processes that constantly regulate posture and movement – processes that function without reflective awareness or the necessity of perceptual monitoring" (Gallagher 2005: 37-38). The second is, as pointed out, "a (sometimes conscious) system of perceptions, attitudes, beliefs and dispositions pertaining to one's own body" (Gallagher 2005: 37). Given these definitions, we can state that acts of proto-mimesis involve (above all) the *body schema*, which is largely innate (in the sense of being present at birth) and pre-conscious, rather than the *body image*, which is gradually constructed with experience and is accessible to consciousness.

Furthermore, the forms of intersubjectivity described here do not require a conceptual differentiation between self and other, which is necessary for establishing a correspondence relation between them, i.e. what we refer in our definition as *representation*. This is not to say that the young infant lives in a completely undifferentiated world in which there is no awareness of self whatsoever, as in classical accounts of infant cognition (e.g. James 1890). Nevertheless, even a modern developmental psychologist who emphasizes the role of awareness of others' attention and the presence of affective self-consciousness in the first months of life points out that "older infants reveal a greater focus on the self and the younger ones reveal a *more immersed, less detached* focus on the other" (Reddy 2003: 401). This "more immersed, less detached" quality of the earliest forms of intersubjectivity supports our classification of them as proto-mimetic.

Can this analysis be extended to the (early) interpersonal relations among apes? As pointed out in Section 4.3 "neonatal mirroring" has also been observed in apes (though this has so far only be attested with human, rather than ape faces as stimuli). Since this is typically attributed to a form of identification with the person imitated, serving as a basis for intersubjectivity when children are concerned (Meltzoff and Gopnik 1993; Gallagher 2005), it can be viewed as evidence that at least apes, and possibly other mammals too, have such basic proto-mimetic intersubjectivity. The function of such "mirroring" can be related to what is possibly the most basic form of intersubjectivity, both ontogenetically and phylogenetically: the ability to share emotions, or *empathy* (*Einfühlung*). As a proto-mimetic, non-representational capacity, this is testified in early infancy and sometimes referred to as *interaffectivity* (Stern 2001). The well-known experiments described by Trevarthen (1992) show that parent-infant interactions in the first few months take the form of a reciprocal rhythmic "dance", and that frustration follows if this "attunement" is disrupted (notice the musical metaphors). The suggestion is that emotions such as joy and suffering are perceived directly, possibly involving mirror-neuron structures similar to those involved in action recognition and imitation, rather than involving inferences to underlying states (Gallese, Keysers and Rizzolati 2004).

Preston and de Waal (2002) have argued persuasively that as a basic mechanism involving the linkage of perception and action, a basic form of empathy is available to most if not all mammal species. Defining empathy as "any process where the attended perception of the object's state generates a state in the subject that is more applicable to the object's state or situation than to the subject's own prior state or situation" (ibid: 4), they see a clear evolutionary motivation for its emergence in the ability to recognize and understand the behavior of con-specifics. It is an open question how much of such matching between the visually perceived body of the other and the proprioceptively perceived body of oneself is domain-general – and thus can be expected to be general across species – and how much is specialized in the form of species-specific communicative signals such as facial expressions. It is characteristic that such signals, such as the famous "play-face" expression of great apes (an evolutionary precursor to the human smile) typically carry emotional rather than referential meaning.

A second socio-cognitive skill that relates to intersubjectivity and appears to have a proto-mimetic origin, at least in human children, is attention. Reddy (2001, 2003, 2005) has argued that prior to awareness of the other's attention to an external object and much prior to *joint attention* appearing around 12 months (see below), children "show an awareness of others as attending beings, as well as an awareness of self as an object of others' attention" (Reddy 2003: 357), displayed in phenomena such as eye-contact, intense smiling, coyness, 'calling' vocalizations, showing-off etc. Since awareness of self (at this stage) is largely proprioceptive, while awareness of the attention of others (in seeing children) is mostly based on vision, this satisfies our first criterion for bodily mimesis. Reddy's claim is that such dyadic (though not *dyadic-mimetic*) interactions underpin later developments in intersubjectivity. Evidence for this is the observation that autistic children show difficulties even with such simple interpersonal engagements, "suggesting that whatever is going on in dyadic attentional engagements may indeed be critical, not just as a source of information and experience about attentional behavior, or as a scaffold for the subsequent development of awareness of attention, but also as evidence of awareness of attention" (Reddy 2005: 95).

Until recently it has not been clear whether such awareness of another's attention exists in the interaction between infant apes and their mothers, but in a recent study, Bard et al. (in press) report that the rates of mutual *gaze* between infants and their mothers are virtually the same in 3-month old human children and 3-month old chimpanzees; 18-20 and 17 times per hour, respectively (though humans tend to engage in longer bouts of mutual gaze). Furthermore, the authors noticed a "cultural" difference between the apes at Primate Research Institue, Japan and those at Yerkes National Primate Research Center, USA, with the ones in Japan engaging in mutual gaze at much higher rates (22 vs. 12 times per hour), while the ape mothers in the USA *cradled* their infants more often (71% vs. 40% of the total time). Intriguingly, a similar inverse corrlelation between visual and tactile contact has been observed in human societies, with traditional cultures favouring touch and Western ones gaze: "With reduced physical contact found in Western societies, mutual engagement shifts to the visual system, arguably a more evolutionarily derived pattern" (Bard et al. in press). Since apes do not seem to differ from humans in the capacity to perform this shift (and possibly even transmit it culturally to their descendents), this confirms the conclusions from neonatal imitation that there is no major difference between the species on the proto-mimetic level.

5.2 Dyadic mimetic intersubjectivity

In our previous discussion, we proposed that proto-mimetic intersubjectivity, without a clear differentiation between self and other is based on the (mostly) unattended body schema and similarly unattended mechanisms for "body copying". On the other hand, what we here refer to as *dyadic mimetic intersubjectivity* is based on the conscious control of the movements of one's body and attention to their correspondence to the body of another, whereby one can *imagine* what the other experiences on the basis of one's own experiences in similar circumstances. In the terminology of Gallagher (1995, 2005), we propose that the role that bodily mimesis proper plays for the development of intersubjectivity implies not the body schema but the *body image*. While the body schema and the body image normally interact, Gallagher (2005) shows how in certain pathologies they can be disassociated.

The distinction between the dyadic form of mimetic intersubjectivity, described here, and the *triadic* one described in the following sub-section is that while the dyadic form involves the properties of volition and representation (see Section 3), e.g. in the case of recognizing oneself in a mirror or in shared attention, there is no understanding of *communicative sign function*. Thus there is no basis for intentional communication on the basis of shared representations. As we shall see here and in the discussion of ape gestures in Section 6, it is the latter that appears to be most difficult for non-human primates. Below, we briefly describe how understanding others' emotions, attention and intentions can be seen as intimately related to dyadic mimesis.

Whereas (simple) empathy is proto-mimetic, what Preston and de Waal (2002) call *cognitive empathy* requires a differentiation between subject and object where "the subject is thought to use perspective-taking processes to imagine or project into the place of the object" (ibid: 18). Evidence that this is not an isolated phenomenon, but shows a more advanced level of intentionality is the fact that cognitive empathy "appears to emerge developmentally and phylogenetically with other 'markers of mind' … including perspective taking …, mirror self-recognition …, deception, and tool-use." (ibid: 18). Research concerning cognitive empathy in apes has focused on their consolation behavior, which is well attested in at least chimpanzees, but has not been found in monkey species (de Waal and Aureli 1996) or any other mammalian species. Consolation is cognitively more complex than simple empathy since the consoling individual not only feels that somebody else experiences a negative emotion, but also intends to help relieve this, implying an ability to imagine the more positive emotional state.[2]

This supports the interpretation that cognitive empathy involves a more sophisticated representational capacity than what is necessary for simple empathy. Since dyadic mimesis involves both the ability to identify with the other, and at the same time to differentiate between self and other, our (Vygotskyan) hypothesis is that it is dyadic mimesis, implicated in e.g. imitation (Section 4) that scaffolds the development of such representational capacity.

Since dyadic mimesis allows to "place oneself in the shoes of others", it also gives the opportunity to understand what someone else is *attending* to. Such "second-order attention" is well testified among great apes (Hare et al. 2000). When two individuals become aware that the other is attending to the same object, what results is *shared attention*. This comes a good deal towards the construction of a "consensual reality" that can be communicated about, but does not quite reach it. To make a given object X fully intersubjective between you and me, I would need not only to "see that you see X", (second-order attention, see Figure 1a), but also "to see that you see that I see X" (third-order attention, see Figure 1b) and vice versa – which is our interpretation of what it means to engage in *joint attention*.

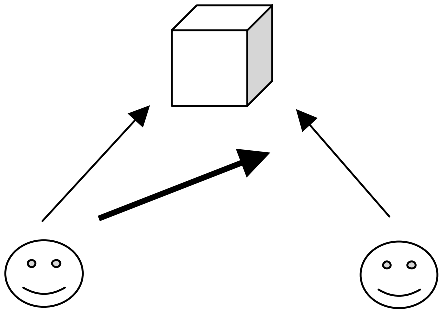

Figure 1a. *Shared attention*: Second-order attention: "I see that you see X" (and vice versa)

---

[2] An interesting question is what goes on in the mind of an ape that is being consoled: If one could identify cognitive processes of the form "I can feel that you feel that I feel bad", this would be equivalent to "third order emotion". However, given the lack of methodology for testing the existence of such a process, we feel complelled to apply Occam's razor and ascribe only second order emotions of the form "I feel that you feel".

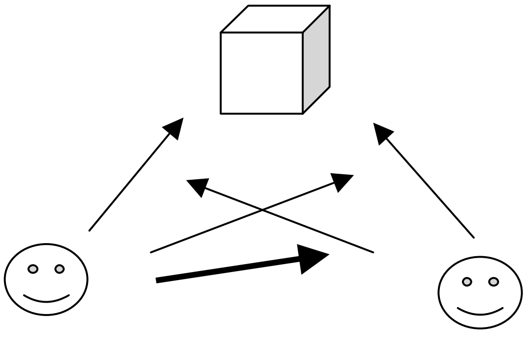

Figure 1b. *Joint attention*: Third-order attention: "I see that you see that I see X" (and vice versa)

Full joint attention thus involves third-order mentality and (possibly because of this, see below) appears to be beyond the cognitive capacities of apes (Tomasello 1999). It also goes beyond dyadic mimesis, so we will leave this capacity for the time being, but return to it in the following subsection, since there is a close connection between communicative intention and the *emergence* of joint attention.

Concerning the understanding of another's intentions, it has been for sometime a predominant opinion that apes cannot do this, and indeed this was one of the central claims of Tomasello (1999). Recently, however, there has been mounting evidence that (at least) "[c]himpanzees understand psychological states – the question is which ones and to what extent", which is the title of Tomasello, Call and Hare (2003). In the experiments behind this claim a subordinate and a dominant chimpanzee competed for food placed on the subordinate's side of two barriers, so that in some cases only the subordinate, but not the dominant could see the food. The subordinates could also monitor the visual access of their dominant competitor. The results showed that the subordinates preferentially retrieved the food that dominants could not see, suggesting that chimpanzees understand what con-specifics have and have not seen and even have some awareness of the dominant's *goal* (i.e. "a mental representation of a desired state", Tomasello et al. in press) of obtaining the food. This does not have to be done on the basis of explicit reasoning or inference, but possibly through the "projection" of one's own motivational state in a similar situation onto the other ("I would get the food if I were in his place!"), in a dyadic mimetic fashion. Apparently, even non-enculturated and language-naïve apes are capable of this.[3]

However, it has not been shown that (non-enculturated) apes are capable of understanding another's mental states about their own mental states, which would involve, again, third-order mentality.[4] As suggested earlier, joint attention also requires third-order mentality, and is similarly something that apes seem to find difficult.

In natural settings, some cases of deception *may* be interpreted in a way to involve third-order mentality, but do not require this. For example de Waal (1998 [1982]) describes the chimpanzee Yeroen who has had a fight with Nikkie, and continues to fake a limp only when in the presence of Nikkie, apparently in an act of wishing to provoke Nikkie's empathy: "I wish to make you see that I hurt". The more parsimonious explanation, however, is that Yeroen has learned from previous experience that he is not bothered by Nikkie when he is hurt: "He may have learned from incidents in the past in which he is seriously wounded that his rival was less hard on him during periods when he was (of necessity) limping" (ibid: 35-36), so he *mimes* the appropriate behavior. Here we have a clear correspondence between dyadic mimesis and second-order intersubjectivity. Notice that Yeroen's limping was not a form of intentional communication – he did not wish Nikkie to understand that he was faking a limp – if he did, that would be a clear case of triadic mimesis.

Another possible instance of third-order mentality is that reported by Woodruff and Premack (1979) in which chimpanzees learned to suppress glances and body orientation toward the place where food was hidden, when a trainer who would not share the food was present. However, since it took a long number of trials

---

[3] While it still not conclusively shown that apes are capable of such mental "projection", and it is conceivable how the evidence can be explained in a more behavioristic manner involving learning generalizations over other individuals' behavior in relation to food, we contend that the mental explanation is (a) ultimately more parsimonious and (b) consistent with other mental skills in apes, as well as in human children in roughly comparable stages of development.

[4] Note that when we in this section link triadic mimesis and third-order mentality, we do not imply that an agent capable of triadic mimesis can handle *all forms* of third-order mentality, in particular those which involve *beliefs*.

for the apes to learn this behavior, a simpler explanation involving conditioning is possible.

The conclusion that we can draw from these various examples is that wild apes as well as those who are exposed to different degrees of human contact (captive, nursery-raised and laboratory-trained apes), but are not raised in a “something like a human cultural environment” and thus *enculturated* (Call and Tomasello 1996) can indeed understand second-order mentality. However, such apes do not seem able to understand third-order mentality – neither in the domains of emotion, attention nor intention – in which their own mental state needs to be either intentionally communicated – in collaboration, or hidden – in competition. This corresponds well with the capacity of apes for dyadic mimesis, but their relative difficulty with triadic mimesis, which we discuss further below.

5.3 Triadic mimetic intersubjectivity

In the case of triadic mimesis, there is by definition not only an understanding of representation itself, but an understanding that representations can be used communicatively, or what we refer to as *communicative sign function.* This implies an understanding that the sign has *the same meaning for the addressee as for the sender.* This involves at least second-order mentality, which was shown above to correlate with dyadic mimesis. But “having the same meaning” is a reflexive notion and this implies at least some degree of third-order mentality. Consider the simple example of what knowing the meaning of the word *cat* implies:

a) I know that *cat* means “a small furry animal that meows”.
b) I assume you know that *cat* means “a small furry animal that meows”.
c) I assume that you know that I know that *cat* means “a small furry animal that meows”.

While it is possible for intentional communication to begin without a full realization of (c), it is practically inevitable that discursive experience (including failures in communication) will promote the development of third-order mentality. Thus, it is possible that it was *sign use itself* that was the major driving force behind the development of intersubjectivity in hominid evolution. Unlike competing hypotheses related to “Machiavellian intelligence” (Byrne and Whiten 1988), this puts the focus on *cooperation* rather than competition (see also Brinck and Gärdenfors 2003; Tomasello et al. in press). A prediction from this hypothesis is that enculturated apes – and these have all been taught at least some communication through signs – will develop higher-level skills of intersubjectivity. There is some support for this prediction. In summarizing some 200 studies of the role of human influence, Call and Tomasello (1996) conclude that “[t]he sociocognitive domains in which humans seem to have the highest effect on apes are intentional communication and social learning” (ibid: 391).

In Section 6, we will focus more specifically on the understanding and use of communicative gestures by apes. Here and in the following subsection, we will mostly review the evidence that sign use in general, and the acquisition of (aspects of) language in particular is the major factor behind the development of higher forms of intersubjectivity that may deserve the label “theory of mind”.

As pointed out earlier, wild apes do not seem to be capable of engaging in full, third order joint attention. Furthermore, as Tomasello (1999: 21) points out, wild apes (1) do not *point to* objects; (2) do not *hold up* objects to show them to others; (3) do not *take someone along* to a place to show them something; (4) do not actively *offer* something to someone; and (5) do not intentionally *teach* other individuals new behaviors. Tomasello’s original account of these absences was based on the claim that apes are unable to understand other’s intentions. Given the more recent evidence, this explanation is no longer tenable, and indeed Tomasello et al. (in press), suggest instead that the crucial difference between apes and humans involves the motivation to participate in joint collaborative engagements, and the lack of this motivation prevents apes from constructing “dialogical cognitive representations”.

Our explanation is similar but more specific: we believe that non-enculturated apes fail to develop the communicative sign function, and related to that, the ability to engage in *third-order* mentality. The motivational difference between apes and humans appealed to by Tomasello et al. (in press) cannot be the full explanation since enculturated apes such as Koko, Kanzi and Chantek manage at least (1) and apparently communicative skills (2-5) above as well (Miles 1999), even if in restricted forms. This seems to imply that the human cultural environments[5] of the enculturated apes have taught them the basics of intentional, sign-mediated communication, and thereby (the roots of) third-order mentality.

How this could occur can be seen again with respect to joint attention, which can be seen as emerging from second-order attention combined with the recognition of another's intention concerning my attention: "I see that you see X" (second-order attention, Figure 1a), and furthermore "I realize that you want me to look at X" (see Brinck 2001). In other words, joint attention can be brought forth by understanding a simple form of communicative intention, combined with already exisiting second-order attention. Thus, communicating the intention to jointly attend may be said to involve the simplest kind of triadic mimesis: whatever kind of behavior that is used to convey that intention – some form of index (see the example in the next paragraph) – can be said to *stand for* that intention for both sender and interpreter.

Without enculturation, experiments indicate that apes do not understand communicative intentions. A rather typical example is an experiment by Tomasello, Call and Gluckman (1997), where the authors in different ways indicated for both chimpanzees and two- to three-year-old children which out of three containers contained a reward: by pointing to the correct container; by placing a marker on top of the correct container; and holding up a replica of the correct container. Tomasello (1999: 102) summarizes the results of the experiment as follows:

> Children already knew about pointing, but they did not know about using markers and replicas as communicative signs. They nevertheless used these novel signs very effectively to find the reward. In contrast, no ape was able to do this for any of the *communicative signs* that they did not know before the experiment. One explanation of these results is that the *apes were not able to understand that the human beings had intentions toward their own attentional states*. The apes thus treated the communicative attempts of the human as discriminative cues on par with all other types of discriminative cues that have to be laboriously learned over repeated experiences. The children, in contrast, treated each communicative attempt as an expression of the adult's intention to direct their attention in ways relevant to the current situation. [our emphasis]

In other words, while the children clearly understood the communicative intentions of the experimenter, the apes did not. This interpretation is supported by a similar experiment designed to test "false beliefs" (Call and Tomasello 1999), in which the enculturated and language-taught orangutan Chantek clearly performed differently from all the other apes in understanding a human communicator's signals. Even though this was not the goal of the experiment, and Chantek did not score better than the other apes in the false beliefs task, his much better performance could be explained by considering that he understood the signals as communicative signs (in this case indexes), rather than as "discriminative cues".

Finally, we should mention the case of captive apes living in a zoo and thus involved with at least some degree of interaction with human culture. In their study of spontaneous gestural communication in a group of gorillas in the San Francisco zoo, Tanner and Byrne (1996, 1999), found a wealth of gestures used by several members of the group, in particular by the adult male Kubie. We return to this is Section 6, but it suffices here to note that the gestures seemed to be used in a communicative way so that:

> [w]hether the receiving partner was a human or another ape, the signaling ape made sure that visual contact was established (except for tactile close gestures), and seemed to *understand both the other's potential actions and what the partner might, in turn, understand from his (the signaler's) performance of gestures.* (Tanner and Byrne 1999: 231, [our emphasis])

We can conclude that a triadic from of mimetic intersubjectivity, involving understanding not only con-specifics' intentions, but their communicative intentions, and consequently a degree of third-order mentality, appears to be *not completely beyond the cognitive potential of apes*. To realize this potential, they need an environment that is rich in opportunities for developing the communicative sign function, i.e. a particular form of enculturation. Thus triadic mimesis may be said to be within apes' "Zone of Proximal Development"

---

[5] Consisting of what Wittgenstein (1953) called the "forms of life" that provide the necessary context for the emergence and functioning of intentional communication and language.

(ZPD) – the notion introduced by Vygotsky (1978) to refer to skills that children could acquire with the help of adults, but not alone.

If "enculturation" provides the ZPD for present-day apes, it is reasonable to suppose that it did the same for some particularly social group of hominids through a form of "self-domestication" giving rise to a bootstrapping spiral of sign use and intersubjectivity. In the terms of Donald (2001), triadic mimesis must have been within the common ancestor's "zone of proximal evolution".

5.4 Post-mimetic intersubjectivity

What differentiates post-mimetic, or symbolic, cognition from mimesis is the use of *fully conventional* signs, interrelated within a *system* (Deacon 1997; Zlatev 2003). The most obvious example of post-mimesis, involving all the previous features but also symbolicity is a conventional, institutionalized signed language such as ASL (Stokoe 1960) or Swedish Sign Language (Ahlgren 2003). What is the relation between acquiring such a system and intersubjectivity?

A convention (Lewis 1969; Clark 1996) or a norm (Itkonen 1978) exists as a form of *mutual* (or *common*) *knowledge* among the members of the group that share the convention. A common explication of mutual knowledge is that it consists of *third-order* knowledge: "I know that you know that I know X". Translated to the capacity to understand the minds of others, common knowledge amounts not only to my understanding the mentality of others, but also to my understanding that others (can) understand my mentality. This has been expressed by Itkonen (1978: 96) as follows:

> [P]erson A cannot know that he is doing or thinking X, and thus cannot do or think X, unless he is able to know what it is for some other person B to do and think X. The same is true, in turn of B's knowledge of his own actions with respect to A's (possible) actions and thoughts. Hence it can be shown on purely conceptual grounds that A must be able, in principle, to identify B's various mental states or processes, and vice versa. This means that mental states and processes exist only at the level of *common knowledge* i.e. of a common ability to identify them, wherever they occur.

Mutual knowledge is often stated as a form of "third-order belief", but this is potentially misleading since it is not necessary that the understanding on all three orders of mentality is explicit enough to be a matter of "belief", i.e. a propositional representation that is actively *held* to be true. Consider again the three orders of knowing the conventional meaning of *cat* mentioned earlier: I assume that you know that I know that *cat* means "a small furry animal that meows". The highest order, my assumption that you know that I know, is not properly speaking a *belief* for the 4 year old child, since it is taken for granted, without pondering on whether it is true or not. At the same time, children at this age become (a) capable of understanding that others lack knowledge or have "false beliefs" (e.g., Perner 1991; Wimmer et al. 1988; Gopnik and Graf 1988; Mitchell 1997), implying *metarepresentational capacity*, and (b) at least somewhat fluent in their first language. It appears that these two developments are closely connected, and that acquiring a language, spoken or signed, is a major causal factor for developing a fully-fledged "theory of mind". Despite successes in understanding other's attention and even intentions, apes, including enculturated ones, have so far consistently failed so-called "false belief tests" (Premack 1988; Call and Tomasello 1996, 1999; Povinelli 2000). They have also failed in acquiring a full human language – despite successes in acquiring certain aspects of it. Three different sides to language (use) combine to promote metarepresentational capacity.

First, as mentioned above, language is a conventional symbolic system, and as such its mastery implies third-order knowledge, which would carry with it training in the understanding of others' beliefs. Second, two specific (universal) features of human languages are (a) mental predicates such as "think", "believe", "know"… and (b) sentential complement constructions such as "say that…". If one can *meaningfully* formulate sentences such as "I think that you think that X", then one should be able to think the corresponding *thought*. Third, as pointed out by Tomasello (1999), not just the logical structure of language, but its use in discourse would promote the understanding of others as "mental agents": There are at least "three kinds of discourse, each of which requires [children] to take the perspective of another person in a way that goes beyond the perspective-taking inherent in comprehending individual linguistic symbols and constructions." (ibid: 173): disagreements, repairs/explanations and meta-discourse.

Empirical evidence for the existence of a strong connection from language to the understanding of beliefs has been accumulating during recent years and involves e.g. the following:

- Deaf children who are not exposed to sign language at an early age pass false belief tasks significantly later than those who are (Peterson and Siegal 1995);
- There are correlations between parent use of mental predicates in their child-directed speech and the children's performance in false belief tasks (de Villiers and Pyers 1997);
- Longitudinal studies indicate that language development predicts false belief task performance, but not vice versa (Astington and Jenkins 1999);
- Training in (non-mental predicate) sentential complement constructions significantly improves performance on false belief tasks (Hale and Tager-Flusberg 2003; Lohmann and Tomasello 2003).
- Training in perspectival discourse alone (without sentential complements or mental predicates) contributes to performance on false belief tasks (Lohmann and Tomasello 2003).

Our conclusion is therefore that the understanding of (false) beliefs is a form of post-mimetic intersubjectivity, in the sense that is based on language, either spoken or signed. Notice that this does *not* contradict analyses by e.g. Bloom (2000) that the acquisition of language presupposes "theory of mind" skills, since we have argued that the latter, including joint attention and communicative intention are triadic mimetic phenomena, which also in our analysis are prerequisites for language.

In one of the relevant studies, Call and Tomasello (1999) used a non-verbal false belief task with chimpanzees and orangutans as well as with human children. The main results were that the children's performance on verbal and nonverbal false belief tasks was highly correlated, supporting the hypothesis of a possible causal connection. At the same time, no ape could pass the nonverbal false belief task even though they succeeded in all of the control trials indicating mastery of the general task demands. However, as mentioned in Section 5.3, in this latter experiment the enculturated and language-taught orangutan Chantek performed differently from the other apes, displaying a better understanding of the experimenter's signals as communicative signs. A prediction from our analysis would be that if Chantek, or any of the other "langauge apes" that have been the subject of so much controversy, could progress in their language development to involve more mental terms, as well as grammatical constructions involving sentence complementation, they would also be able to pass false belief tasks.

A counter-example to this Vygotskyan hypothesis of interpersonal communication preceding intrapersonal cognition would be to show that non-enculturated apes are capable of understanding beliefs (or at least third-order mentality), even outside of communicative settings. A candidate domain is again deception, but as mentioned in Section 5.2, the kind of examples involving deception, usually provided in support for theory-of-mind skills (e.g. Yeroen's fake limp) can be explained by simpler skills. A further candidate is *counter-deception.* In a study performed with chimpanzees by Menzel (1973, 1974), only the female Belle was allowed to see where food was hidden within a large enclosed area, and subsequently she led the group to the food, and all were able to share it. When the experiment was repeated, however, the male Rock began to take all the food for himself. After this, Belle did not reveal where the food was hidden if Rock was nearby. She sat further and further from the place where the food was hidden, and it was only when Rock was looking in another direction that she went to the hiding place. Rock, however, countered this by pretending to go away and then turning just as Belle was about to reveal where the food was. Byrne (1995) interprets this as implying that Rock has an understanding that Belle is trying to deceive him, i.e. third-order mentality. However, Tomasello and Call (1997) note that the experiment with all its steps lasted several months so Belle and Rock may have successively learned how to predict the behaviour of the other without understanding deceptive intentions. This hypothesis is supported by a later experiment with mangabey monkeys performed by Coussi-Korbel (1994), who replicated Menzel's study, but performed a detailed day-to-day analysis. One of the informed mangabeys, step by step in four days, learned the strategy of leading her main competitor away from the food.

5.5 Summary

In this section we have tried to show that there is a close connection between different levels of intersubjectivity and the different levels of the mimesis hierarchy, as summarized in Table 4 below.

Table 4. The mimesis hierarchy and intersubjectivity

Level	*Intersubjectivity skills*	*Type of mentality*
Proto-mimesis	(simple) empathy mutual attention	1st order: lack of complete differentiation between self and other
Dyadic mimesis	Cognitive empathy Shared attention Understanding other's intention (in competitive contexts)	2nd order: understanding the other through "projection" (identification, but differentiation)
Triadic mimesis	Joint attention Communicative intentions	3rd order (attention and intentions)
Post-mimesis	(False) belief understanding	3rd order (beliefs)

Proto-mimesis is crucially implicated in mutual attention and the awareness of others' feelings, through a species-general capacity for empathy that has possibly been further developed in the "ultra-social" species *Homo sapiens*. Dyadic mimesis leads to the ability to map between one's own body and that of others, in a more detached, differentiated way, and in this way understand others' emotions, i.e. cognitive empathy, shared attention and even intentions through a conscious process of "projection": what would I see/feel/wish if I were you. Unlike earlier claims to the contrary, newer evidence and analyses show that apes do not have much difficulties with this level and that they have the capacity for second-order mentality.

The crucial step in the evolution of intersubjectivity involves triadic mimesis, implying having and understanding others' communicative intentions, which requires third-order mentality: "I want you to do X (e.g. share attention on an object) by recognizing my intention that you do this" from the sender's perspective and "I understand that you want me to do X" from the recepient's.

This is not something that comes easily to apes in natural conditions, but through enculturation and especially through extensive sign use, some understanding of communicative intentions seems to be within the reach of apes' "Zone of Proximal Development", even though in may be in its periphery. Evidence for this is the relative mastery of joint attention by enculturants, and as argued this can be seen to originate in (dyadic mimetic) second-order attention combined with the understanding of the other's intention that I attend. Finally, post-mimesis, with its features of conventionality and systematicity necessarily brings with it sufficient understanding of third-order mental states, and (at least) second-order beliefs, e.g. "I think that you know (or don't)". The latter have been shown in numerous false belief tasks, which are regularly failed by apes. On the other hand, there is mounting evidence that the acquisition of language, with its structural properties and discursive use significantly contributes to higher forms of "theory of mind" development.

In summary, in this section we have argued that bodily mimesis, in its proto, dyadic and triadic forms, is a major factor in the development of intersubjectivity. Sign use itself was suggested to be a driving force in the development of an understanding of third-order mentality, and the performance of enculturants shows that this achievement is within the reach of apes. Therefore one can conclude that (mimetic) sign use was possibly within the "zone of proximal evolution" (Donald 2001) of the common ape-human ancestor, considerably more so than fully-fledged language, characterised by full conventionality and systematicity.

Finally, while we have focused on the mimesis-to-intersubjectivity direction of causality, this does not mean that there is no "feedback" link and thus we assume a co-evolutionary scenario. It was implicit in the discussion that achieving different levels of intersubjectivity is necessary for moving to still higher levels on the mimesis hierarchy, e.g. joint attention for the acquisition of language. It is also possible that in the present analysis we have overemphasized the Vygotskyan principle of the "interpsychological preceding the intrapsychological" (see Section 1), but this has come rather naturally by defining the developmental levels or stages in terms of bodily mimesis, with its originally interpersonal character.

6. Mimesis and gesture

Of the the three socio-cognitive domains we discuss in this article, gesture is the one most obviously related to bodily mimesis: at least some gestures posses all the properties emphasized by Donald (1991): reference, intentionality, communicativity, autocuing and generativity (see Section 3). However, not all bodily acts that are called "gestures" have these properties. Furthermore, in the currently highly active field of *gesture studies* (e.g. Kendon 1981; McNeil 1992; Goldin-Meadow 1999; Kita and Özyürek 2003), studying among other things the spontaneous co-speech gesturing of human children and adults, it is often pointed out that gestures and language interact in complex ways. Thus, not all gestures are easily subsumed under the label "non-linguistic" or "non-verbal". In such a case, it becomes questionable whether they are truly mimetic rather than post-mimetic.

Perhaps in the most general sense gesture has been defined as "a functional unit, an equivalence class of coordinated movements that achieve some end" (Armstrong, Stockoe and Wilcox 1995: 46; Wilcox 2004: 44), which encompasses communicative as well as non-communicative body movements as long as they posses the property of goal-directedness. This is much too general for our purposes and we will consider as gestures only *goal-directed communicative body movements*, *i.e. such that require interpretation from an audience for achieving the gesturer's goal.* This includes a variety of acts such as attention getting, pointing and miming. At the same time, our definition excludes bodily signals such as laughter, which is not goal-directed, or acts that achieve their ends through "mechanical causality" rather than interpretation, e.g. pulling someone, as opposed to beckoning him. Notice that while we focus on communication, we do not require all gestures to be *intentionally* communicative: what is essential is that the bodily motion is interpreted by an observer as revealing something about the state of mind of the gesturer.

There are numerous classifications of gestures (see Kendon 1981 for a well-known taxonomy), but none, to our knowledge, that has explicitly taken an evolutionary perspective – apart from the analysis of the origins of pointing, see below. Our goal in this section is to apply the mimesis hierarchy to gesture, and in showing how different types of gestures correspond to the different stages of the mimesis hierarchy, suggest such an evolutionary taxonomy. We will also show how the four stages of the hierarchy cross-cut with two basic semiotic distinctions: *indexicality* and *iconicity*, which differ with respect to the *semiotic ground* (Peirce 1931-1935; Sonesson 1989) for the relationship between the "sign vehicle" (the expression) and the meaning of the gesture:

- Indexical gestures, where the ground is one of spatio-temporal *contiguity*, the prototypical case being *pointing.*
- Iconic gestures, where the ground is of *similarity*, with the prototype *pantomime.*

Figure 2 illustrates the two types schematically: what is essential for indexical gestures, as with all indexical signs, is that expression and content are closely related in time-space. An iconic gesture (sign), on the other hand, requires similarity (often of a quite schematic type) but on the other hand it can be removed in time-space from their meanings.

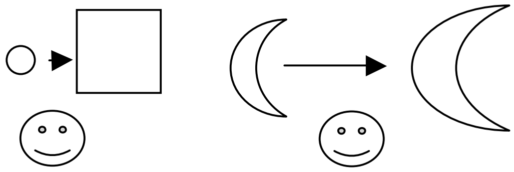

Figure 2. Indexical signs, based on *contiguity* vs. iconic signs, based on *similarity*. In both case expression and content need to be differentiated and perceived as standing in a representational relationship by an interpreter for there to be a (genuine) sign. Without this, there are only indexicalities and iconicities.

In agreement with Sonesson (1989, in press), we distinguish between iconicity and indexicality as a type of semiotic ground, and the *sign function* (Piaget 1945). The latter presupposes an interpreter (signified by the "happy face" in Figure 2) who differentiates between the expression and the content and is furthermore aware of the representational relationship between them: Only in this case do we have iconic and indexical *signs*. In the case of gestures, this means that only if the gesturing individual is aware of the expression-content structure of his gesture are the gestures intentionally used as signs. If not, the gestures are only a matter of indexicality or iconicity for the gesturer; however, they may still be seen as signs by an *interpreter*, as is the case with the first types of gestures discussed below.

6.1 Proto-mimetic gestures

We have defined proto-mimesis as involving a cross-modal mapping between exteroperception and proprioception, but lacking (clear) differentiation and or/volition. From this perspective the spontaneous reaching behavior of the child observed soon after birth (von Hofsten 1983) and various forms of phylogenetic ritualization may be classed as proto-mimetic (for the gesturer), involving indexicality and iconicity, respectively, but not the sign function.

6.1.1 Proto-mimetic indexicality: reaching

In the first 6 months of life, the human child coordinates its hand movements and reaches towards interesting objects, including towards such which it desires but cannot grasp because they are too far away. At the same time there is no gaze alternation between the object and a third party, and no attempt to get the other's attention – and possibly help – in obtaining the object. Clearly this is not a communicative gesture from the point of view of the child. However, at least in typical Western caregiver-child interactions at this period of development, the adult often interprets such reaching as communicative and offers the desired object. This appears to have an effect on the child's interpretation of its own action through a process of internalization of the interpersonal to intrapersonal (Vygotsky 1978, see Section 1). Eventually, this transforms reaching into (proto-) imperative pointing, which, as discussed below, is a dyadic mimetic gesture.

Ape infants reach out to objects that attract their attention similarly to human children, but less so to distal objects due to the fact that they become capable of locomotion much sooner. Furthermore, ape mothers are less "interventionist" and less likely to oblige their infants, even when the latter do reach out to distal objects (Bard and Vauclair 1984). In general they interfere only when they deem that there is a risk of their offspring getting in harm's way. Such differences may be the major reason why feral apes point much less (if at all) than human beings, since they are quite capable of acquiring pointing gestures when in contact with humans (see below).

6.1.2 Proto-mimetic iconicity: phylogenetic ritualization

Phylogenetic ritualization is often proposed as the basic mechanism for the emergence of animal signals, as explained by e.g. Knight (2002: 147):

> Over evolutionary time, certain aspects of non-communicative behaviour assume a signalling function, becoming correspondingly specialized through a process known as 'ritualization'. Signal evolution begins when others read some aspect of normal behaviour as significant. If the subject of such surveillance can benefit from having its mind read or its behaviour anticipated, then over evolutionary time natural selection will accentuate the cues, reducing any ambiguity. By definition, such phylogenetic ritualization entails special elaboration and added costs.

Gestures such as gorilla chest beating (and other forms of display) and the "submissive gesture" of apes, in which they crouch and "make themselves small" involve *iconicities* (resemblances) to body size – the larger the animal is, the louder the signal – and state of agitation – the stronger the emotion, the more beating. At the same time, they are not iconic *signs*, since the similarity between the signal and the non-communicative behaviour is not perceived (at least not necessarily) by either the gesturer or the interpreter. Hence, while there is communication, there is no representation (in the sense of expression-content structure, see Ikegami and Zlatev in press) involved. That is why we can call such phylogenetically ritualized signals proto-mimetic gestures.

It is characteristic that when observed in highly enculturalted apes such as the signed language taught gorilla Koko, gestures such as "gimme" and "pound" were the first to develop, with the least effort on the part of the trainer. As pointed out by Patterson (1980: 517):

> These were distinctive because they appeared without direct training at a time when the great majority of Koko's signs were being acquired by molding. Because other young gorillas reared in captive conditions who have not been exposed to sign language use these same gestures, they must be part of the gorilla's natural repertoire.

6.2 Dyadic mimetic gestures

In the case of dyadic mimesis there is *representation* in the sense that there is both differentiation and correspondence between the gesture and its "meaning", in this case the *action* that the gesturer wishes that the addressee will perform. In this respect, such gestures constitute signs, in the sense of the term defined at the beginning of this section. However, they have two major limitations: (a) their intended meaning is always a desired action rather than a statement, i.e. they are imperative rather than declarative, and (b) they involve second-order mentality: a claim on the attention and behavior of the addresse, but not third-order mentality: there is no appeal on the addresse to understand the gesturer's own mentality. In our terminology, they do not involve the communicative sign function, unlike the triadic gestures that we discuss in 6.3.

6.2.1 Dyadic-mimetic indexicality: imperative pointing

In the human infant literature "reaching" is often used as a label for acts of intentional communication, rather than to attempts att prehesion. Similarly to Leavens and Hopkins (1999), in their informative review of studies of pointing in children and apes, we find that even if the child does not shape the hand and arm in a "canonical point" (Butterworth 1998), given that there is evidence for gaze alternation between an object and a third party, the act should be regarded as an instance of pointing. In human children, such acts emerge sponatenously from 8 months of age, and at least at first have an "imperative" function of *attempting to make the addressee perform some desired action.* From this it follows that imperative pointing requires second-order mentality, but not necessarily joint attention and third-order mentality (see Section 5.2). Hence, imperative pointing can be classified as a dyadic mimetic gesture. This conclusion is also supported by Tomasello's (1999) observation that in this stage of development, children can master intentional pointing without understanding the pointing of others. Another kind of pointing that is perhaps even more obviously dyadic mimetic that is also observed at this stage is "pointing for oneself", i.e. when the child is using pointing gestures to direct its own attention, rather than that of others (Bates, Camaioni and Volterra 1975).

When it comes to apes, it was earlier thought that apes do not point (Donald 1991) or at least not with the index finger (Povinelli and Davis 1994). This has been recently shown to be untrue. Leavens and Hopkins (1999) summarize some of the evidence showing that pointing has been reported for 15 captive monkeys, 83 captive apes, and even one feral ape, a bonobo who was pointing at a group of humans while looking back and forth between the humans and the other bonobos (Veà and Sabater-Pi 1998).[6] Whether or not feral apes really point, the evidence from *captive* apes (e.g. Call and Tomasello 1994; Leavens, Hopkins and Bard 1996, Krause and Fouts 1997; Leavens and Hopkins 1999) clearly shows that the criteria for imperative pointing are satisfied: attention getting, gaze alternation and persistence until the goal is met. Leavens and Hopkins calls these *audience effects*, and observe that captive chimpanzees engaged in imperative pointing exhibit more audience effects, and above all more gaze alternation than children at the corresponding stage of development. At the same time, the children were considerably more prone to vocalize while pointing, showing the close association between bodily mimesis and speech in human beings.

However, if we limit our attention to the manual gestures alone we can conclude that at least when exposed to human contact, apes clearly display imperative pointing, requiring dyadic mimesis. Simarly, there is evidence that at least enculturated apes engage in dyadic, non-communicative pointing in examining objects for the purpose of directing their own attention

---

[6] If it was not for the somewhat anecdotical character of this piece of evidence, it would pose a problem for our classification, since it appears to involve *declarative* pointing, which is otherwise observed only in enculturated apes.

(Bonvillian and Patterson 1999). Interestingly such non-communicatieve pointing emerged earlier than communicative pointing (of both the imperative and the declarative type) in the sign language taught gorilla Koko, similar to the case with children (Bates, Camaioni and Volterra 1975).

6.2.2. Dyadic mimetic iconicity: ontogenetic ritualization

As pointed out in 6.1, most animal signals are the result of phylogenetic ritualization, and are in this sense innate. But monkeys and especially great apes use a good number of learned gestures as well. Tomasello et al. (1994) distinguished six functional categories of gestures in infant and juvenile chimpanzees involving nursing, walk-riding, grooming, eating, appeasing and play/tickle. These were believed to have come about through a process called *ontogenetic ritualization* in which the gestures were shaped by responding to each other's anticipations. This happens over the life of the individual rather than the species, hence the term given to this learning process. Tomasello and Zuberbühler (2002: 295) describe ontogenetic ritualization as a sequence of 4 steps:

> (1) Individual A performs behavior X.
> (2) Individual B reacts consistently with behavior Y.
> (3) Subsequently B anticipates A's performance of X, on the basis of its initial step by performing Y.
> (4) Subsequently A anticipates B's anticipation and produces the initial step in a ritualized form (waiting for a response) *in order to* illicit Y. [original emphasis]

This way of describing the emergence of such gestures sounds rather behavioristic, implying that it is only a matter of (co-)reinforcement of behavioral patterns. However, Tomasello et al. (1994) showed that most chimpanzee gestures involve getting the other's attention and that the gestures vary with the attentional state of the recipient: non-tactile, non-vocal gestures are predominantly used when there is visual contact between the recipient and the gesturer. Furthermore, the animals clearly waited for a response after gesturing. The presence of such "audience effects" indicates that there was an *intention* on the side of the gesturing chimpanzee to produce an effect on the addressee. However, this intention to communicate is not the same as a *communicative intention*, which is an intention that the desired effect will be produced by the addressee *recognizing* the intention to produce the effect (Grice 1957; Sperber and Wilson 1995). As argued in the Section 5, such forms of third-order mentality appear to be beyond the capacity of non-enculturated apes.

At the same time, to the extent that the gesturing ape is differentiating between its (ontogentically ritualized) gesture and the desired effect and being more or less aware of the correspondence relation between them, it is justified to regard the two as standing in a representational relationship – even without involving a communicative sign function. However, this representational relationship is fairly unstable. Tomasello et al. (1994) showed that chimpanzes had a limited arsenal of gestures that they applied highly "flexibly": different gestures were used for the same goal, and the same gesture was often used with different "meanings". On the one hand, this speaks for these gestures being intentional and different from species-specific signals. But on the other hand, it testifies to their being non-conventional, and non-systematic.

Similar findings have been replicated for young gorillas (Pika, Liebal and Tomasello 2003) and bonobos (Pika, Liebal and Tomasello 2005). However, in these more resent (and systematic) studies, a number of gestures where found to be group-specific, and therefore "some form of social learning" (Pika et al. 2003: 108), besides ontogenetic ritualization was considered to be the mechanism behind their emergence. Furthermore, while Tomasello et al. (1994) did not observe any social transmission of a specific begging gesture learned by a chimpanzee to acquire rewards from human experimenters, Pika et al. (2005) summarize a number of different studies finding evidence for the social learning of gestures in great apes. It is therefore not possible to exclude some degree of imitation in apes' learning of gestures. Even so, that does not imply that such gestures, e.g. the "arm shake" of the gorillas of Apenheul zoo (Pika et al. 2003) are more than dyadic mimetic, since it is not clear that they involve the communicative sign function: e.g. *represent something for someone.*

Some dyadic mimetic gestures may also involve iconicity (similarity) between the spatio-temporal profile of the gesture and the meaning of the gesture: the action to be performed by the addressee. An example would be to gently rub down someone's back in the desired direction of movement instead of forcefully pushing the other down. A purely visual version would be to make the gesture in the air in view of the recipient. Such gestures are usually labeled *iconic* (Tanner and Byrne 1996, 1999) and we will follow this practice. However, we still need to ask whether they function as iconic *signs* for both the sender and the receiver, i.e. the iconicity is grasped as the ground for the representational relationship. What characterises dyadic mimesis, as we stated above, is that this relationship is (a) unstable and (b) non-symmetrical: it is possible that only one of the communication parts perceives the iconicity as an iconic sign.

In a study of captive bonobos by Savage-Rumbaugh et al. (1977) almost all gestures that preceded copulation appeared to be signs of copulatory positions. Since bonobos use many different copulatory positions it is expected that some mutual agreement must be reached about which position to be used. Two animals could sometimes "argue", in the sense that one gestured for a ventro-ventral position while the other gestured for a dorso-ventral one. The likelihood of unusual copulation postures increased if the bout was preceded by gesturing. These gestures were grouped into three groups: (1) positioning motions: the recipient's limbs are manipulated by the initiator. The actual movement is done by the recipient but the direction is set by the initiator; (2) touch plus iconic hand motions: a touch on the part of the body that should be moved, followed by an indication with the hand how this part should move; (3) iconic hand motions: without touching the other individual the gesturer shows what he wants the other to do; these are the most abstract ones.

It is possible to interpret this progression as the development of an understanding of iconicity and an emergence of iconic signs (Savage-Rumbaugh et al. 1977), but it should be noted that the gestures from the first group were the most frequent in the bonobo group, and gestures from the most abstract level, the third group, was the least frequent. In other studies of bonobos (de Waal 1988; Pika et al. 2005) and gorillas (Pika et al. 2003) no gestures could be readily interpreted as iconic in nature. So even if the evidence seems to grant apes some understanding of iconic signs, predominantly emerging through ontogenetic ritualization, this understanding does not seem to be reflective enough to be systematically used in communication.

6.3 Triadic mimetic gestures

As pointed out repreatedly above, what is lacking in dyadic mimetic gestures is an understanding of the communicative sign function, which implies that the gesturer not only wishes the addressee to perform some desired action, but to do this by appreciating the representational (sign) relation between his gesture and its referential meaning. As we show in this section apes can achieve such an understanding, but clearly only when they have been taught aspects of a conventional symbolic system, i.e. a language. This raises some doubts whether such gestures are not actually post-mimetic.

6.3.1 Triadic mimetic indexicality: declarative pointing

The third stage in the development of human children's pointing, mastered around 14 months, is usually called *declarative pointing* (Bates, Camaioni and Voltera 1975; Brinck 2004). The crucial difference with respect to imperative pointing is that the child need not desire the object pointed, but rather the act of achieving joint attention with the addressee on the object is the goal in itself. As argued in Section 5.3, joint attention involves understanding a simple form of communicative intention, and is thus triadic mimetic. In this sense declarative pointing is also triadic mimetic.

Brinck (2004) proposes that at an early stage of a child's use, an additional purpose of declarative pointing is the *evaluation* of the indicated object, achieved by an exchange of emotional information about it. For example, the child points to an object in order to obtain information from the addressee whether the object is dangerous. The main benefit for the child of this kind of exchange is that he/she can learn about objects vicariously. This primary function presupposes that the child can understand the emotions of the addressee, and as we argued in Sections 5.1 and 5.2, this is well within the competence of the child at this stage of development, and similarly for apes. Thus, it is possible that the evaluative side of declarative pointing was the original cause of its proliferation among the hominids.

Enculturation has been shown to have dramatic effects on the pointing of apes. While feral apes almost never point, as mentioned earlier, captive apes and monkeys readily learn to engage in imperative, though not declarative pointing. On the other hand, at least three of the documented language-taught apes – Koko, Chantek and Kanzi – consistently point declaratively, i.e. refer to objects and locations for sake of calling these to the attention of someone else (Patterson 1980; Miles 1990). In the mentioned review of pointing in human children and apes, Leavins and Hopkins (1999), point out that language-trained apes were more likely to use a canonical point with the index finger than children of the age 13-18 months. In some studies, it is even suggested that language-taught apes combine such an act of declarative pointing and another sign to make a *predication*, i.e. THIS is X (Greenfield and Savage-Rumbaugh 1990).

However, the question arises if such acts are cases of triadic mimesis, rather than post-mimetic skill. The close relation between declarative pointing and reference makes it likely that the process of being taught language itself is what has simultaneously taught the significance of declarative pointing to the apes. The following episode recorded by Savage-Rumbaugh and Levin (1994), involving not the famous, highly enculturated bonobo Kanzi, but the somewhat more traditionally trained Austin and Sherman, shows how enculturation, and specifically the teaching of communicative sign function, can affect pointing:

> Upon spying the new food, Sherman reached out for it and I suggested that he should ask for it. He quickly turned, scrutinized the keyboard, and deliberately, but with no fanfare, indicated one of the unassigned lexigrams as the symbol for the food. Austin watched attentively and subsequently used the symbol Sherman had selected to indicate that he would like to try the new food also. This unexpected skill encouraged us to believe that Sherman and Austin understood that a unique correspondence existed between each food item and a specific symbol. Moreover, they sometimes pointed back and forth between the symbol and the food item. They had named new foods and apparently agreed on coordinated use of the selected symbols. (ibid: 79)

We will return to this issue in the summary below.

6.3.2 Triadic-mimetic iconicity: communicative intention?

Possibly the clearest evidence for iconic triadic mimetic gestures in non-language-trained apes comes from a study of zoo-living gorillas, partially reared by human caregivers. Tanner and Byrne (1996, 1999) studied a pair of western lowland gorillas, Kubie and Zura, at the San Francisco Zoo, who seemed to have an unusually rich repertoire of gestures: audible, silent and tactile ones. The gestures were used almost exclusively in play sessions and sexual situations, and the choice of gesture differed according to the attentional state of Zura, with audible and tactile ones used when she was not looking, which mirrors the findings for chimpanzees by Tomasello et al. (1994).

Some of Kubie's gestures were "attention getters" and play invitations. His most successful gesture was an armswing under his belly towards his genitals. It had the highest frequency of accompanying "play faces" and resulting in physical contact. It was only performed when the gorillas had visual contact. "Armswing under" was often preceded by a tap on Zura's body. This is comparable with Savage-Rumbaugh et al.'s (1977) second group of iconic gestures: touch plus iconic hand motions. If Zura failed to respond, a gesture was often repeated, strengthened or more gestures were made – this also resembling the pattern found in chimpanzees (Tomasello et al. 1994). Most of the gestures ended in contact between Kubie and Zura within 5 seconds, with Zura making most of the contact, suggesting that she was (correctly) interpreting Kubie's gestures. In 66% of the cases when a tactile gesture was made Zura moved her body in the direction indicated by Kubie. If a gesture was coupled with a "play face" the likelihood of physical contact was significantly higher than for gesture alone. However, "head nod", Kubie's most frequent gesture, had a high degree of play bouts without using "play faces", which shows that Zura did respond to gestures and not just to the natural (pre-mimetic) gorilla signals.

Still, the question remains if Zura really responded to their (intended) meaning, comprehending the gestures as iconic signs. Associating playfulness (or sexual agendas) with particular behaviors in another individual's repertoire is not sufficient to show that Zura understood Kubies gestures as communicative signs. Whether or not Kubie's gestures had an iconic relationship to their referents for him, they could in each case remain indexical to Zura, i.e. be behavioral correlates to moods. The iconicity of some of Kubie's gestures could very well have been abstractions from more physical manipulations shaped through ontogenetic ritualisations (Tomasello and Call 1997) and Zura's appropriate response be derived from the context.

However, over time Kubie used his gestures with several females, and if he did not go through an identical abstraction phase with all of them, which of course cannot be ruled out completely, how could the receiver understand them? Tanner and Byrne (1996) suggest that comprehension of motion depicted in gesture may be biologically encoded for the great apes since their adaptation to brachiation requires a three-dimensional understanding of space, and space is the medium for gestures.[7]

In summary, like the bonobo sexual gestures described in 6.2.2, the San Francisco gorilla gestures could be argued to be at least dyadic-mimetic, since they correspond to actions intended to be performed by the recipient. But through their iconicity the gestures also have a potential to be triadic-mimetic, under the condition that at least the recipient, and possibly also the sender understood their sign function. Triadic mimesis cannot be excluded as an alternative to ontogenetic ritualization in the formation and use of such gestures because non-conventional gestures could be understood not only through the context but through the iconicity of the gestures themselves. If, and it is a big "if", Kubie, Zura and the other females could use the iconic gestures effectively because of such referential understanding, Kubie's gestures would be an example of triadic mimesis. More empirical research is necessary to decide this issue.

A final point should be made in relation to the differences observed between feral and enculturaed apes pointed out earlier. The fact that Zura, the other female gorillas and Kubie all at some point had been reared by humans (Tanner and Byrne 1996) might have implications for the conclusion we can draw about mimetic capacities in non-enculturated apes. Human beings engage readily in triadic forms of communication and apes with extensive exposure to human language show good comprehension of other representational mediums, like words and pictures (Savage-Rumbough, Shanker and Taylor 1998). However, it is not yet known how much exposure to human culture is needed to yield an understanding of triadc mimesis (if at all) in hand-reared zoo animals, and if this would be retained into adulthood.

While the presence of iconic signs in captive apes remains unsure, there is solid evidence for their use in enculturated apes. The sign-language taught gorilla Koko reportedly invented a number of such signs herself, including signs referring to *actions* e.g. "blow", "tickle", "walk-my-back", "bite", as well as *objects*: "bird", "bracelet", "stethoscope". These gestures clearly involve a referential triangle involving the gesture, referent and the interpreter, and were sometimes even used in the absence of the referent, showing one of the crucial features of language: displacement (Hockett 1960), and even "metaphorically", e.g. using "bird" as an insult.

A problem for classifying them as triadic mimetic, however, is that these gestures were quite similar to the actual signed language signs that she was taught. In accordance with the caveat stated earlier, it is difficult to exclude the interpretation that they are actually post-mimetic, i.e. induced by the language-teaching situation itself.

Some support for their mimetic-iconic nature is offered by the fact that Koko produced many so-called gestural *modulations* with respect to size, manner, number, etc. – all having an iconic basis. For example, the signs for a "small bird" and a "big bird" would differ in size; those for "several birds" would involve repetition of the sign. Thus, gestural modulations are perhaps the clearest evidence of the presence of triadic mimesis in Koko's productions, especially since Patterson (1980: 539) emphasizes that they "were created by Koko herself". Other testimony for the spontaneous production of triadic mimetic gestures in a language-trained ape can be found in the spontaneous gestures of the bonobo Kanzi, who unlike Koko was not taught signed language but a combination of visual lexigrams (in production) and spoken English (in

---

[7] Brachiation also entails an ability that is unique to apes: that of being able to move the wrist, elbow and shoulder joints in all directions, making subtle gestures possible.

comprehension). Greenfield and Savage-Rumbaugh (1990, 1991) report on Kanzi's spontaneous use of pointing (see below), and at least some gestures involving iconicity such as "come", "go" and "chase".

6.4 Post-mimetic gestures

Mimesis was defined by Donald (1991) to be "non-linguistic", and in Section 3 and in our definition of bodily mimesis we clarified what we take this to mean: lack of full conventionality (mutual knowledge) of the signs and systematicity, involving at least a degree of compositionality. To the extent that the gestures of apes display these properties, they can be regarded as post-mimetic.

Evidence from four of the most successful projects involving the teaching of language to great apes – the chimpanzee (*Pan troglodytes*) Washoe (Fouts 1972, 1973), the gorillas Koko and Michael (Patterson 1978, 1980) and the bonobos (*Pan paniscus*) Kanzi and Panbanisha (Savage-Rumbaugh and Lewin 1994; Savage-Rumbaugh et al. 1998) and the orangutan Chantek (Miles 1990) – has shown that many of the essential characteristics of language are within the grasp of our nearest non-human relatives. As is the case with children a precondition for the success of these projects has been a cultural environment rich with intersubjectivity and a variety of activities to stimulate communication (Miles 1990). Our conclusion is similar to that of Miles (1999: 204), namely that all great apes "have the intelligence for a rudimentary, referential, generalizable, imitative, displaceable, symbol system".

The reason that this claim has been controversial (e.g. Terrace et al. 1981) is not foremost empirical, but rather a matter of a disagreement on what the "essential characteristics" of human language are. Turning to the data reported in the "ape language" literature we see convincing evidence that apes can:

- comprehend the referential (representational) function of spoken words, ASL signs and visual lexigrams, and combinations of these;
- use the sign-tokens in the absence of their referents, i.e. "displacement" (Hockett 1960);
- acquire a considerable vocabulary of words/signs, according to some measurements extending 600 signs, but even according to the most conservative criteria no less than 140 signs;
- regard the acquired signs as conventional-normative (consensual), to the point of correcting their teachers if the latter do not use these appropriately;
- understand novel combinations of spoken or signed words;
- produce novel combinations of signs;
- use language for a number of different functions (speech acts), including labeling, answering, expressing emotion, arguing and insulting;
- use language not only for communication, but for thinking (private speech).

What has not yet been clearly demonstrated is whether the spoken or signed utterances of apes conform to consistent principles of grammatical organization, though e.g. Greenfield and Savage-Rumbaugh (1990, 1991) describe at least two fairly consistent ordering rules in the two symbol combinations of Kanzi, one of them involving a pointing gesture (see below).

It is thus unsurprising that those who have (over-)emphasized syntax as opposed to semantics as the defining characteristic of language (e.g. Bickerton 1990; Pinker 1994; Hauser, Chomsky and Fitch 2002) continue to deny "the proposition that language is no longer the exclusive domain of man" (Patterson 1978: 95). But even if the language of apes such as Koko and Kanzi is termed "proto-language" rather than true language, it is still qualitatively different from mimesis, since it possesses at least rudiments of conventionality and systematicity. In this sense, some, though not all, of the gestures of the language-taught apes can be regarded as post-mimetic.

6.4.1 Post-mimetic indexicality: paralinguistic pointing

At least two forms of the pointing of language trained apes seem to go beyond mimesis, since the pointing gesture is integrated in the (proto-) linguistic system acquired. First of all are the combinations of a pointing gesture referring to an object or an individual with another sign expressing a predicate applying to the individual, thus forming a *predication*. Furthermore, the bonobo Kanzi has been reported to form such "predications" – with imperative as well as declarative functions – in a fairly systematic manner Greenfield and Savage-Rumbaugh (1990, 1991).

A related form post-mimetic indexicality involves *labeling*, in which the pointing gesture directed at an object, or a picture of one in a book, is accompanied with the the production of (gestural) sign denoting the object. Such clearly declarative uses of pointing + label, some of which are performed "privately", i.e. not for the benefit of a human interlocutor (Bodamer et al. 1994) are perhaps the most convincing case for the presence of post-mimetic, paralinguistic pointing in enculturated apes. Kanzi has been shown to use pointing + lexigram, as well as lexigram + pointing combinations for the sake of labeling quite frequently: altogether there were 249 examples in his five-month corpus of two-element combinations in which Kanzi indicated an entity gesturally and named it with a lexigram (Greenfield and Savage-Rumbaugh 1991).

6.4.2 Post-mimetic iconic gestures: Signed language

Human signed languages offer the clearest picture of post-mimesis. Recent studies of the spontaneous emergence of Nicaraguan Sign Language (NSL) during the past 25 years show that signed languages have their origin in (triadic) mimesis, but quickly acquire the properties of conventionality and systematicity. Senghas, Kita and Özyürek (2004) compared the co-speech gestures of Nicaraguan speakers of Spanish, with the signing of three "cohorts", or generations, of learners of NSL and could document some aspects of this transition in detail:

> The movements of the hands and body in the sign language are clearly derived from a gestural source. Nonetheless, the analyses reveal a qualitative difference between gesturing and signing. In gesture, manner and path were integrated by expressing them simultaneously and holistically, the way they occur in the motion [event] itself. Despite this analogue, holistic nature of the gesturing that surrounded them, the first cohort of children, who started building NSL in the late 1970s, evidently introduced the possibility of dissecting out manner and path and assembling them into a sequence of elemental units. As second and third cohorts learned the language in the mid 1980s and 1990s, they rapidly made this segmented, sequenced construction the preferred means of expressing motion events. NSL thus quickly acquired the discrete, combinatorial nature that is hallmark of language.

Given their mimetic-gestural origin signed languages have a much greater degree of iconicity than spoken languages and it has been proposed that this plays a role in their faster acquisition by (deaf) children (Brown 1977). Recent studies have questioned this, however, since only a minority of the signs of signed language have transparent iconic meanings, and in a study of 22 children acquiring ASL it was shown that "of the 44 different signs produced by the children before the age of 13 months, 36% were classified as iconic, 30% as metonymic, and 34% as arbitrary" (Bonvillian and Patterson 1999: 253).[8] In this study, the authors compared the rate and pattern of acquisition of ASL by children and that of two gorillas, Koko and Michael. Despite certain differences – the children's acquisition was (unsurprisingly) faster – it was shown that "that the similarities in early development across the species outweigh the differences" (ibid: 260). Thus, it can be concluded that gorillas, and by inference other great apes, not only can acquire the basics of a post-mimetic symbolic system such as ASL, but that they do this is a similar way. Interestingly, however, the gorillas seemed to rely somewhat more on the iconicity of the signs in comparison with the children, so that the proportion of their first 46 signs was somewhat different to the one reported above: 42% iconic, 32% metonymic and 26% arbitrary, while for the first 10 signs this difference was even clearer: 60% iconic, 20% metonymic and 20% arbitrary. We interpret this as suggesting that the apes relied to a greater degree on their skills of triadic mimesis than the children in their acquisition of the sign language, which also would explain why the children quickly progressed beyond the initial level of vocabulary acquisition to

---

[8] "Metonymic" signs are such that involve some degree of iconicity between the sign and the referent, but "the tie between the sign and its meaning is not readily apparent – one would be unlikely to guess the meaning of a metonymic sign simply seeing it produced" (Bonvillian and Patterson 1999: 252). In the distinction introduced by Sonesson (2001) they involve *secondary* (rather than primary) iconicity, which can be perceived first after knowing the specific sign relationship.

learn the systematic character, i.e. the grammar, of the language, while the apes "stagnated" on a simple, "proto-language" level.

6.5 Summary

In this section we have reviewed some of the evidence on the production, and to some extent the comprehension of gestures by great apes, and when relevant compared it to the acquisition of gestures and sign languages by children. Applying the mimesis hierarchy to this evidence has resulted in the classifiction summarized in Table 5.

Table 5. The mimesis hierarchy and ape gestures

Level	*Indexical gestures*	*Iconic gestures*
Proto-mimesis	Reaching	Some phylogentically ritualized signals
Dyadic mimesis	Pointing for oneself Imperative pointing	Some ontogenetically ritualized gestures
Triadic mimesis	Declarative pointing	Iconic signs
Post-mimesis	Paralinguistic pointing	(Aspects of) signed language

Some of the natural signals of apes can be regarded as involving indexicality and iconicity, without these being perceived as signs by either the sender or the receiver. On the other hand, if an observer and, especially, a caregiver during infancy *does* see these as signs, this can provide an impetus for making these gestures mimetic, as seems to be the case with the origin of imperative pointing. At least some ontogenetically ritualized gestures appear to involve differentiation and correlation between the gesture and meaning, and thus can be regarded as simple iconic signs. Since these, however, can function without having and interpreting communicative intentions, they are (at most) at the level of dyadic mimesis. Feral and even captive apes have not been convincingly shown to progress beyond this level.

Enculturation, and especially language-teaching changes this and consequently both declarative pointing and iconic gestures emerge, testified in individuals such as the gorilla Koko, the bonobo Kanzi and the orangutang Chantek. We interpret this as evidence for triadic mimesis, though with some hesitation, since it is likely that being taught language was necessary for the apes to acquire the communicative sign function. Even so, triadic mimesis can be shown to be within the grasp of our nearest non-human relatives. In comparison, their skills in post-mimesis involving full conventionality and systematicity, i.e. language, are more limited. The conclusion, similar to that at the end of Section 5, is that triadic mimesis is within the "zone or proximal evolution" of great apes, and by inference, to our common ancestor.

7. Summary and conclusions

The goal of this study has been to explore the potentials of bodily mimesis to provide perhaps not "the" missing link in explaining human cognitive evolution, but at least one important missing piece in the puzzle, and thus help unravel the mystery of the origins of language. In Section 3 we defined the concept of bodily mimesis and the related *mimesis hierarchy*, thereby refining and modifying the concept of mimesis proposed by Donald (1991). Our method has been to aply the mimesis hierarchy to evidence from primatology (and to some extent child development) in order to be able to compare ape skills in three different domains: imitation, intersubjectivity and gesture and to determine whether any general patterns could be discerned. Table 6 below summarizes our results.

Beginning with *proto-mimesis*, we could see that there is abundant evidence from all three domains that apes excel in this. There do not appear to be any (major) cross-species differences with humans in this respect, and we can assume with some confidence that the common ape-human ancestor had comparable skills in imitation, intersubjectivity and communicative gestures.

On the next level, that of *dyadic mimesis*, we begin to discern some differences – both between the levels, and between humans and apes. There is clear evidence for e.g. cognitive empathy and ontogenetically ritualized gestures in apes – and we suggested how these phenomena can be explained in a dyadic-mimetic fashion involving both identification with and differentiation from the other. However, when it comes to imitation, apes appear to have difficulties in performing detailed imitation on the *action level*, which is the type of imitation where bodily mimesis is most relevant. Human children in comparison excel in this and even "overmimic". This difference between the species does not appear to be due to a difference in intersubjectivity, i.e. understanding of others' intentions, as was originally suggested by Tomasello (1999) but rather in mimetic capacity per se. Thus, our results confirm, at least in part, Donald's proposal that (bodily) mimesis constitutes the phenomenon in relation to which we should look for human cognitive specificity. At the same time, the results reviewed in Section 4 showed clearly that apes have the basic capacity for dyadic mimesis, and are thus beyond the limitations imposed by "episodic cognition" (Donald 1991).

In contrast, when we come to the capacities of apes to perform *triadic mimesis* we find predominantly negative evidence in all three domains. Enculturated apes such as Kanzi, Chantek and Koko are capable of this with respect to gesture (e.g. declarative pointing) and intersubjectivity (e.g. joint attention), and possibly also, but less clearly with respect to the learning and/or invention of signs through imitation, which also here appears to be "lagging behind" on the mimesis hierarchy. As for the evidence concerning non-enculturated apes, it is much less certain that they reach the level of triadic mimesis, but best with respect to *iconic gesture* (e.g. Tanner and Byrne 1996, 1999). We can generalize and suggest that apes appear to have difficulties with *third-order mentality*, involving both attention and intention. However, as shown by the studies of enculturated apes, through the practice of communicative sign use, apes have the potential to understand communicative intentions, thereby showing that even triadic mimesis is within apes' "zone of proximal development".

Finally, the capacity of apes for *post-mimesis*, relying on the possession of a conventional symbolic system, is furthest away from the potential of human beings. While enculturated apes such as Kanzi may have mastered a proto-language after a good deal of exposure and effort, this remains limited in grammatical and semantic complexity, approximately to the level of a two-year old child. Thus, even post-mimesis is not completely beyond the ZPD of apes, though it is further in the periphery than triadic mimesis. Using the analogy suggested by Donald (2001), we can say that both triadic mimesis and post-mimesis are within apes' "zone of proximal

Table 6. Comparing ape skills in the domains of imitation, intersubjectivity and communicative gestures along the levels of the mimesis hierarchy: each level mentions example capacities, with "+" indicating positive evidence and "–" negative evidence. W = wild, C = captive, E = enculturated.

Level	**Imitation**	**Intersubjectivity**	**Communicative gesture**
Proto-mimesis	*Neonatal mirroring* C +	*Mutual attention* *Empathy* C +	*Phylogenically ritualized gestures* W/C +
Dyadic mimesis	*Action level imitation* C + (?)	*Shared attention* *Cognitive empathy* C +	*Ontogentically ritualized gestures* *Imperative pointing* C +
Triadic mimesis	*Imitatititave signs* W – C – E + (?)	*Joint attention* W – C – E +	*Declarative pointing* W – C + (?) E +
Post-mimesis	*Role-reversal imitation* –	*Belief understanding* –	*Proto-language* E +

evolution", i.e. the sphere in which evolutionary adaptations are likely to accur, but since post-mimesis presupposes triadic mimesis (even conceptually), the latter constitutes the most likely missing link.

What kind of factors brought this about in evolution? The evidence summarized and analyzed in this article allow us to make two high-level generalizations concerning the three domains of bodily mimesis analysed:

(1) Ape skills of imitation are comparatively *least* developed along the mimesis hierarchy;
(2) Ape skills of gesture are comparatively *most* developed along the mimesis hierarchy.

Furthermore, we could discern a possible causal relationship in the direction from gesture (sign use) toward intersubjectivity: the only apes who succeed in tasks of joint attention, where those who where taught to use sign system through enculturation. (In the case of imitation the relation is more of a logical sort: since triadic imitation is *defined* as the imitation of signs, it is impossible to perform this without having first learned the latter.) Thus we can conclude by suggesting the following two (partially independent) hypotheses of human cognitive evolution, corresponding to the two generalizations given above:

(1´) The crucial adaptation that differentiated early hominds from apes involved *detailed bodily imitation*, allowing for the development of specific *mimetic schemas* which are pre-linguistic representational, intersubjective structures, emerging through imitation but subsequently interiorized (see Piaget 1945; Zlatev 2005a; Zlatev, Persson and Gärdenfors in press).

(2´) Since apes display some potential for spontaneously developing indexical and iconic gestures – e.g. in the rather unusual environment of the San Francisco Zoo, see Tanner (2004) – it is reasonable to assume that the common ape-human ancestor possessed such a potential as well. An "ecological niche" requiring a high degree of intentional communication (there are various scenarios for bringing this about; for a proposal regarding the role of Oldowan material culture see Osvath and Gärdenfors 2005) would have spurred the use of iconic and indexical gestures, which on their part would have both utilized and futher developed mimetic schematic representations.

In sum, our general proposal is that the source of human cognitive specificity lies above all in two, mutually co-enforcing factors involving bodily mimesis: (a) the capacity to form *mimetic schemas*: representations deriving from the imitation of public events and (b) the use of mimetic schemas for intentional communication: *triadic mimesis*. The first of these factors is predominantly representational and the second, communicative, and it is likely that in evolution the two have co-evolved. We submit that this proposal is consistent with the evidence involving the capacities and limitations of apes, and shows the explanatory potential of the concept of bodily mimesis as "a missing link" in our theoretization of cognitive and linguistic evolution.

References

Ahlgren, I. (2003) Teckenspråk. In: *Sveriges officiella minoritetsspråk. Finska, meänkieli, samiska, romani, jiddisch och teckenspråk. En kort presentation.* Nordstedt.

Anisfeld, M. (1996) Only tounge protrusion modeling is matched by neonates. *Developmental Review*, 16: 149-161.

Arbib, M. (2003) The evolving mirror system: A neural basis for language readiness. In: *Language Evolution*, ed. M. Christiansen and S. Kirby, Oxford University Press.

- (in press) From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences.*

Armstrong, D., Stokoe, W., and Wilcox, S. (1995) *Gesture, and the Nature of Language.* Cambidge: Cambridge University Press.

Astington, J. W. and Jenkins, J.M. (1999) A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology,* 35/5: 1311-1320.

Badcock, C. (2000) *Evolutionary Psychology. A Critical Introduction.* Cambridge: Polity Press.

Bard K. A., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A. and Matsuzawa, T. (in press) Group differences in the mutual gaze of chimpamzees. *Developmental Psychology* Vol. 41, 2005.

Bard, K. A. and Vauclair, J. (1984) The communicative context of object manipulation in ape and human adult-infant pairs. *Journal of Human Evolution*, 13: 181-90.

Baron-Cohen, S. (1995) *Mindblindedness: An Essay on Autism and Theory of Mind.* Cambridge, MA: MIT Press.

Bates, E., Camaioni, L. and Volterra, V. (1975) Performatives prior to speech. Merrill-Palmer Quarterly, 21: 205-226.

Bering, J. M., Bjorklund, D. F. and Ragan, P. (2000) Deferred imitation of object-related actions in human-reared juvenile chimpanzees and orangutans. *Developmental Psychobiology,* 36: 218-32.

Bickerton, D. (1990) *Language and Species.* Chicago: University of Chicago Press.

- (2003) Symbol and structure: A comprehensive structure for language evolution. In: Christiansen and Kirby (2003), 77-93.

Bjorklund D. F., Bering, J. M. and Ragan, P. (2000) A two-year longitudinal study of deferred imitation of object manipulation in a juvenile chimpanzee (*Pan troglodytes*) and orangutan (*Pongo pygmaeus*). *Developmental Psychobiology*, 37: 229-37.

Bjorklund, D. F. and Bering, J. M. (2003) A note on the development of deferred imitation in enculturated juvenile chimpanzees (*Pan troglodytes*). *Developmental Review*, 23: 389-412.

Bjorklund, D. F., Yunger, J. L., Being, J. M. and Ragan, P. (2002) The generalization of deferred imitation in enculturated chimpanzees (*Pan troglodytes*). *Animal Cognition*, 5: 49-58.

Bloom, P. (2000) *How Children Learn the Meaning of Words.* Cambridge, MA: MIT Press.

Bodamer, M. D., Fouts, R. S., Fouts, D. H., and Jensvold, K. B. (1994) Indicating acts during counting by a chimpanzee (*Pan troglodytes*). *Journal of Comparative Pscyhology*, 109: 47-51.

Boesch, C. (2003) Is culture a golden barrier between human and chimpanzee? *Evolutionary Anthropology*, 12: 82-91.

Boesch, C. and Tomasello, M. (1998) Chimpanzee and human cultures. *Current Anthropology*, 39/5: 591-614.

Bonvillan, J. D. and Patterson, F. G. (1993) Early sign language acquisition in children and gorillas: Vocabulary content and sign iconicity. *First Language*, 13: 315-38.

- (1999) Early sign-language acquisition: Comparison between children and gorillas. In: *The Mentalities of Gorillas and Orangutans: Comparative Perspectives*, eds, S.T. Parker, R.W.Mitchell, H.L. Miles, 240-264. Cambridge University Press.

Brinck, I. (2001) Attention and the evolution of intentional communication. *Pragmatics and Cognition,* 9/2: 255-272.

- (2004) Joint attention, triangulation and radical interpretation: A problem and its solution, *Dialectica*, 58(2): 179-206.

Brinck, I. and Gärdenfors, P. (2003) Co-operation and communication in apes and humans. *Mind and Language,* 18(5): 484-501.

Brown, R. (1977) Why are signed languages easier to learn than spoken languages? In: *Proceedings of the National Symposium on Sign Language Research and Training*, ed. W. Stokoe, 9-24. Silver Spring, MD: National Association of the Deaf.

Butterworth, G. (1988) What is special about pointing? In: *The Development of Sensory, Motor, and Cognitive Capacities in Early Infancy: From Perception to Cognition*, ed. F. Simion and G. Butterworth 171-190. Hove: Psychology Press.

Byrne, R. W. and Stokes, E. J. (2002) Effects of manual disability on feeding skills in gorillas and chimpanzees. *International Journal of Primatology*, 23/3: 539-54.

Byrne, R. W. (1995) *The Thinking Ape: Evolutionary Origins of Intelligence.* Oxford: Oxford University Press.

- (1998) Commentary on Boesch, C. and Tomasello, M. Chimpanzee and human culture. *Current Anthropology*, 39: 604-5.
- (1999) Imitation without intentionality. Using string parsing to copy the organization of behaviour. *Animal Cognition*, 2: 63-72.

Byrne, R. W. and Russon, A. E. (1998) Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, 21: 667-21.

Call, J. and Tomasello, M. (1994) Production and comprehension of referential pointing by orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, 108: 307-317.

- (1996) The effect of humans in the cognitive development of apes. In: *Reaching into Thought: The Minds of Great Apes.* ed. A. Russon, K. Bard and S. Parker. Cambridge: Cambridge University Press.
- (1998) Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and human children (*Homo sapiens*). *Journal of Comparative Psychology*, 112/2: 192-206.
- (1999) A nonverbal false belief task: The performace of children and great apes. *Child Development,* 70/2: 381-395.

Call, J., Hare, B., Carpenter, M. and Tomasello, M. (2004) "Unwilling" versus "unable": chimpanzees' understanding of human intentional action. *Developmental Science*, 7/4: 488-98.

Cheney, D. and Seyfarth, R. (1990) *How Monkeys See the World: Inside the Mind of Another Species.* Chicago: University of Chicago Press.

Christiansen, M. and Kirby, S. (2003) (eds.) *Language Evolution.* Oxford: Oxford University Press.

Clark, H. (1996) *Using Language.* Cambridge: Cambridge University Press.

Corballis, M. C. (2002) *From Hand to Mouth: The Origins of Language.* Princeton, NJ: Princeton University Press.

Coussi-Kourbel, S. (1994) Learning to outwit a competitor in mangabeys (*Cercocebus torquatus torquatus*). *Journal of Comparative Psychology,* 108: 164-171.

Custance, D. M., Whiten, A. and Bard, K. A. (1995) Can young chimpanzees (Pan troglodytes) imitate arbitrary actions? Hayes and Hayes (1952) revisited. *Behaviour*, 132/11-12: 837-59.

Custance, D. M., Whiten, A., Sambrook, T. and Galdikas, B. (2001) Testing for social learning in the "artificial fruit" processing of wildborn orangutans (*Pongo pygmaeus*), Tanjung Puting, Indonesia. *Animal Cognition*, 4: 305-13.

Custance, D. M., Whiten, A. and Fredman, T. (1999) Social learning of an artificial fruit task in capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 113/1: 13-23.

de Waal, F. (1988) The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour*, 106/3-4: 183-251.

- (1996) *Good Natured: The Origins of Right and Wrong in Humans and Other Animals.* Harvard: Harvard University Press.

- (1998 [1982]) *Chimpanzee Politics.* Harper.

de Waal, F. and Aureli, F. (1996) *Consolation, Reconciliation and a Possible Cognitive Difference Between Macaques and Chimpanzees.* Cambridge: Cambridge University Press.

de Villiers, J., and Pyers, J. (1997) Complementing cognition: The relationship between language and theory of mind. In *Proceedings of the 21st Annual Boston University Conference on Language Development.* Somerville, MA: Cascadillia Press.

Deacon, T. (1997) *The Symbolic Species: The Co-Evolution of Language and the Brain.* New York: Norton.

- (2003) Universal grammar and semiotic constraints. In Christiansen and Kirby (2003), 111-139.

Dennett, D. (1996) *Kinds of Minds. Towards an Understanding of Consciousness.* London: Phoenix.

Dessalles, J.-L. (2004) Language as an isolated niche. Presented at *The Fifth Conference on the Evolution of Language*, Leipzig.

Donald, M. (1991) *Origins of the Modern Mind. Three Stages in the Evolution of Culture and Cognition.* Harvard: Harvard University Press.

- (1998) Mimesis and the Executive Suite: Missing links in language evolution. In *Approaches to the Evolution of Language*, eds., J. Hurford, M. Studdert-Kennedy and C. Knight, 44-67. Cambridge: Cambridge University Press.

- (2001) *A Mind So Rare. The Evolution of Human Consciousness.* New York: Norton.

Edelman, G. (1992) *Bright Air, Brilliant Fire: On the Matter of the Mind.* New York: Basic Books.

Eldridge, N. and Gould. S. J. (1972) Puctuated equilibria: An alternative to phyletic gradualism. In: *Models in Paleobiology*, ed., T. Schopf. San Fransisco: Freeman.

Fouts, R. S. (1972) Use of guidance in teaching sign language to a chimpanzee (*Pan troglodytes*). *Journal of Comparative and Physiological Psychology*, 80/3: 515-22.

- (1973) Acquisition and testing of gestural signs in four young chimpanzees. *Science,* 180: 978-980.

Gallagher, S. (1995) Body schema and intentionality. In: *The Body and the Self,* ed. J. Bermúdez, N. Eilan, and A. Marcel. Cambridge, MA: MIT/Bradford Press.

- (2005) *How the Body Shapes the Mind.* Oxford: Oxford University Press.

- (in press) Phenomenological and experimental contributions to understanding embodied experience. In Ziemke, Zlatev and Frank (in press).

Gallese, V., C. Keyners and Rizzolatti, G. (2004) A unifying view of the basis of social cognition. *Trends in Cognitive Sciences,* 8/9: 396 – 403.

Gallup, G. (1982) Self-awareness and the emergence of mind in primates. *American Journal of Primatology, 2*: 237-248.

Gardner, R. A. and Gardner B.T. (1969) Teaching sign language to a chimpanzee. *Science*, 165: 664-672.

Givón, T. (2002) *Bio-Linguistics. The Santa Barbara Lectures.* Amsterdan: Benjamins.

Goldin-Meadow, S. (1999) The role of gesture in communication and thinking. Trends in Cognitive Sciences, 3/11: 419-429.

Gómez, J. C. (1994) Mutual awareness in primate communication: a Gricean approach, In: *Self-awareness in Animals and Humans*, ed. S. T. Parker, R. W. Mitchell and M. L. Boccia. Cambridge: Cambridge University Press.

Gopnik, A. and Astington, J. W. (1988) Children's understanding of representational change, and its relation to the understanding of false belief and the appearance-reality distinction, *Child Development,* 59: 26-37.

Gopnik, A. and Graf, P. (1988) Knowing how you know: young children's ability to identify and remember the sources of their beliefs. *Child Development,* 59: 1366-1371.

Gopnik, A. and Meltzoff, A. N. (1997) *Words, Thoughts, and Theories*, Cambridge, MA: MIT Press.

Greenfield, P. M. and Savage-Rumbaugh. E. S. (1990) Grammatical combination in Pan pamiscus: Processes of learning and invention in the evolution and development of language, In: *"Language" and Intelligence in Monkeys and Apes*, eds. S T. Parker and K. R. Gibson, 540-578. Cambridge: Cambridge University Press.

Greenfield, P. M. and Savage-Rumbaugh, E. S. (1991) Imitation, grammatical development, and the invention of protogrammar. In N. Krasnegor, D. Rumbaugh, M. Studdert-Kennedy, and R. Schiefelbusch (Eds.), *Biological and behaviorial determinants of language development* (pp. 235-258). Hillsdale, NJ: Erlbaum.

Greenfield, P. M., and Savage-Rumbaugh, E. S. (1993) Comparing communicative competence in child and chimp: The pragmatics of repetition. *Journal of Child Language*, *20*, 1-26.

Grice, P. (1957). Meaning. *Philosophical Review, 66*: 377-88.

Gärdenfors, P. (2003) *How Homo Became Sapiens: On the Evolution of Thinking*. Oxford: Oxford University Press.

Hale, C. M. and Tager-Flusberg, H. (2003) The influence of language on theory of mind: A training study. *Developmental Science* 6/3: 346-359.

Hare, B., Call, J., Agnetta, B. and Tomasello, M. (2000) Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59: 771-85.

Hare, B., Call, J. and Tomasello, M. (2001) Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61: 139-51.

Hauser, M. (1996) *The Evolution of Communication*. Cambridge, MA: MIT Press.

Hauser, M., Chomsky, N. and Fitch, T. (2002) The faculty of language: What is it, who has it and how did it evolve? *Science*, 298: 1559-1569.

Heyes, C. (1998) Theory of mind in nonhuman primates. *Behavioral and Brain Sciences,* 21: 101-115.

Hockett, C. F. (1960) The origin of speech. *Scientific American,* 23: 89-96.

Itkonen, E. (1978) *Grammatical Theory and Metascience*. Amsterdam: Benjamins.

Ikegami, T. and Zlatev, J. (in press) From pre-representational cognition to language. In Ziemke, Zlatev and Frank (in press).

James, W. (1890) *The Principles of Psychology*. New York: Henry Holt.

Johansson, S. (2005) *Origins of Language. Constraints on Hypotheses*. Amsterdam: Benjamins.

Kendon, A. (1981) Geography of gesture. *Semiotica,* 37: 129-163.

Kita, S. and Özyürek, A. (2003) What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language,* 48: 16-32.

Knight, C. (2002) Language and revolutionary consciousness. In: *The Transition to Language*, ed. A. Wray, 138-160. Oxford: Oxford University Press.

Krause, M and Fouts, R. (1997) Chimpanzee (*Pan troglodytes*) pointing: Hand shapes, accuracy and the role of gaze. *Journal of Comparative Psychology*, 111 (4): 330-336.

Laakso, A. (1993) Language equals mimesis plus speech. Commentary to Donald: *Origins of Mind. Behavioral and Brain Sciences,* 14/4: 765-766.

Landau, M. (1991) *Narratives of Human Evolution*. New Haven: Yale University Press.

Lewis, D. K. (1969) *Convention: A Philosophical Study*. Cambridge MA: Harvard University Press.

Leavens, D. A., Hopkins, W. D. and Bard, K. (1996) Indexical and referential pointing in chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, 110: 346-353.

Leavens, D. A. and Hopkins, W. D. (1999) The whole-hand point: The structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology*, 113/4: 417-425.

Lohmann, H. and Tomasello, M. (2003) The role of language in the development of false belief understanding: A training study. *Child Development* 74/4: 1130-1144.

Liberman, A. M. and Mattingly, I. G. (1985) The motor theory of speech perception revised. *Cognition*, 21: 1-36.

Mandler, J. (2004) *The Foundations of Mind: Origins of Conceptual Thought*. Oxford: Oxford University Press.

Marler, P. (1985) Representational vocal signals of primates. In: *Experimental Behavioral Ecology and Sociobiology*, ed. B. Hölldobler and M. Lindauer. Stuttgart: Fischer Verlag.

\- (1991) Difference in behavioral development in closely related species: Birdsong. In: *The Development and Intergration of Behavior*, ed. P. Bateson, 41-70. Cambridge: Cambridge University Press.

Matsuzawa, T., Biro, D., Humle, T. Inoue-Nakamura, N., Tanook, R. and Yamakoshi G. (2001) Emergence of culture in wild chimpanzees: Education by master-apprenticeship. In: *Primate Origins of Human Cognition and Behavior*, ed. T. Matzuzawa. Springer-Verlag.

McNeill, D. (1992) *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.

Meltzoff, A. and Gopnik, A. (1993) The role of imitation in understanding other persons and developing a theory of mind. In: *Understanding other Minds: Perspectives from Autism*, ed. S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen. Oxford: Oxford University Press.

Meltzoff, A. and Moore, M. K. (1977) Imitation of facial and manual gestures by human neonates. *Science* 198: 75-78.

\- (1983) Newborn infants imitate adult facial gestures. *Child Development,* 54: 702-709.

\- (1994) Imitation, memory, and the representation of persons. *Infant Behavior and Development* 17: 83-99.

Menzel, E. W. (1973) Chimpanzee spatial memory organization, *Science* 182: 943-945.

\- (1974) A group of chimpanzees in a 1-acre field: leadership and communication, In: *Behavior of Nonhuman Primates*, ed. A. M. Schrier and F. Stollnitz. Academic.

Merleau-Ponty, M. (1962 [1945]) *Phenomenology of Perception*. Routledge and Kegan Paul.

\- (2003) *La Nature: Notes, Cours du College de France*. Trans. Robert Vallier, Northwestern University Press.

Miles, H. L. (1990) The cognitive foundations for reference in a signing orangutang. In: *"Language" and Intelligence in Monkeys and Apes*, eds. S T.

Parker and K. R. Gibson. Cambridge University Press.

- (1999) Symbolic communication with and by great apes. In: *The Mentalities of Gorillas and Orangutans*, ed. S. Taylor Parker, R.W. Mitchell and H. L. Miles. Cambridge University Press.

Miles, H. L., Mitchell, R. W. and Harper, S. E. (1996) Simon says: The development of imitation in an enculturated orangutan. In: *Reaching into Thought: The Minds of the Great Apes*, ed. A. E. Russon, K. A. Bard. and S. T. Parker. Cambridge University Press.

Mitchell, P. (1997) *Introduction to Theory of Mind: Children, Autism and Apes*. Arnold.

Mitchell, R. and Miles, H. L. (1993) Apes have mimetic culture. Commentary to Donald: *Origins of Mind. Behavioral and Brain Sciences* 14/4: 768.

Myowa-Yamakoshi, M. (2001) Evolutionary foundation and development of imitation. In: *Primate Origins of Human Cognition and Behavior*, ed. T. Matsuzawa. Springer-Verlag.

Myowa-Yamakoshi, M. and Matsuzawa, T. (1999) Factors influencing imitation of manipulatory actions in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 113/2: 128-36.

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M. and Matsuzawa, T. (2004) Imitation in neonatal chimpanzees (*Pan troglodytes*). *Developmental Science*, 7/4: 437-42.

Nagell, K., Olguin, R. S. and Tomasello, M. (1993) Processes of social learning in the tool use of chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, 107/2: 174-86.

Nelson, K. (1996) *Language in Cognitive Development. The Emergence of the Mediated Mind.* Cambridge University Press.

Nelson, K. and Shaw, L. K. (2002) Developing a socially shared symbolic system. In: *Language, Literacy and Cognitive Development,* ed. J. Byrnes and E. Amseli, Lawrence Erlbaum.

Osvath, M. and Gärdenfors, P. (2005) Oldowan culture and the evolution of anticipatory cognition. *Lund University Cognitive Studies 122*, Lund.

Patterson, F. (1978) The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language*, 5: 72-97.

- (1980) Innovative use of language in a gorilla: A case study. In: *Children's Language, Vol 2*, ed., K. Nelson 497-561. New York: Garnder Press.

Perner, J. (1991) *Understanding the Representational Mind.* Cambridge, MA: MIT Press.

Perner, J., Leekam, S. and Wimmer, H. (1987) Three-year-old's difficulty with false belief: the case for a conceptual deficit, *British Journal of Developmental Psychology,* 5: 125-137.

Peterson, C. C. and Siegal, M. (1995) Deafness, conversation and the theory of mind. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 36: 459-474.

Piaget, J. (1945) *La Formation du Symbole Chez l'enfant.* Delachaux et Niestlé.

- (1953) *The Origin of Intelligence in the Child.* Routledge.
- (1954) *The Construction of Reality in the Child.* Basic Books.

Pierce, C. S. (1931-35) *The Collected Papers of Charles Sanders Peirce*. Vols. 1-4. Cambridge, Mass.: Harvard University Press.

Pika, S., Liebal, K. and Tomasello, M. (2003) Gestural communication in young gorillas (*Gorilla gorilla*): Gestural repertoir, learning, and use. *American Journal of Primatology*, 60: 95-111.

- (2005) Gestural communication in subadult bonobos (*Pan paniscus*): Repetoire and use. *American Journal of Primatology*, 65: 39-61.

Pinker, S. (1994) *The Language Instinct.* William Morrow.

Povinelli, D. J. (2000) *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works*. Oxford University Press.

Povinelli, D. J. and Davis, D. R. (1994) Differences between chimanzees (*Pan troglodytes*) and humans (*Homo sapiens*) in the resting state of the index finger: Implications for pointing. *Journal of Comparative Psychology*, 108: 134-139.

Premack, D. (1988) "Does the chimpanzee have a theory of mind?" revisted. In *Machiavellian Intelligence. Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*, ed. R.W. Byrne and A. Whiten, 160-179. Oxford: Clarendon Press.

Preston, S. and de Waal, F. (2002) Empathy: Its ultimate and proximal bases, *Behavioral and Brain Sciences,* 25: 1-72.

Reddy, V. (2001) Teasing, joking and mucking about in the first year. In: *Natural Theories of Mind*, ed. A. Whiten. Blackwell.

- (2003) On being an object of attention: Implications of self-other-consciousness. *Trends in Cognitive Science*, 7/9: 397-402.
- (2005) Before the "third element": Understanding attention to self. In Naomi Klein (ed.) *Joint Attention*, 85-109. Oxford: Oxford University Press.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996) Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131-141.

Rizzolatti, G. and Arbib, M. (1998) Language within our grasp. *Trends in Neurosciences*, 21: 188-194.

Robinson, P. (1996) *Consciousness, Rules and Instructed Second Language Acquisition.* Peter Lang.

Russon, A. E, and Galdikas, B. M. F. (1993) Imitation in free-ranging rehabilitant orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, 107/2 : 147-61.

Saussure, F. de (1916) *Cours de Linguistique Générale* [Course in General Linguistics]. Payot.

Savage-Rumbaugh, S. E., Wilkerson, B. J. and Bakeman, R. (1977) Spontaneous gestural communication among conspecifics in the pygmy chimpanzee (*Pan paniscus*). In: *Progress in Ape Research*, ed. G. H. Bourne. Oxford: Academic.

Savage-Rumbough, S. and Lewin. R. (1994) *Kanzi: The Ape at the Brink of the Human Mind.* New York: John Wiley.

Savage-Rumbaugh, S., Shanker, S., and Taylor, T. (1998). *Apes, Language and the Human Mind.* Oxford: Oxford University Press.

Searle, J. (1992) *The Rediscovery of the Mind.* MIT Press.

Senghas, A., Kita, S. and Özyürek, A. (2004) Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305: 1779-1782.

Sinha, C. (1988) *Language and Representation. A Socio-naturalistic Approach to Human Development.* Harverster Press.

- (2003) The evolution of language: From signals to symbols to system. In: *Evolution of Communication systems: A Comparative Approach*, ed. K. Oller and U. Griebel. MIT Press.

Sonesson, G. (1989) *Pictorial Concepts.* Lund University Press.

- (2001) From semiosis to ecology. On the theory of iconicity and its consequences for the ontology of the Lifeworld. *VISIO* 6/2-3: 85-110.
- (*in press*) The "Symbolic species" revisited. Considerations on the semiotic turn in cognitive science and biology.

Sperber, D. and Wilson, D. (1995) *Relevance: Communication and Cognition.* Cambridge University Press.

Stern, D. (2000) *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology.* Basic Books.

Stoinski, T. S. and Whiten, A. (2003) Social learning by orangutans (*Pongo abelii* and *Pongo pygmaeus*). *Journal of Comparative Psychology*, 117/3: 272-82.

Stoinski, T. S., Wrate, J. L., Ure, N. and Whiten, A. (2001) Imitative learning by captive western lowland gorillas (*Gorilla gorilla gorilla*) in a simulated food-processing task. *Journal of Comparative Psychology*, 115/3: 272-81.

Stokes, E. J. and Byrne, R. W. (2001) Cognitive capacities for behavioural flexibility in wild chimpanzees (*Pan troglodytes*): the effect of snare injury on complex manual food processing. *Animal Cognition*, 4: 11-28.

Stokoe, W. (1960) Sign language structure: An outline of the visual communication system of the American deaf. *Studies in Linguistics, Occasional Paper 8,* University of Buffalo, New York.

- (2001). *Language in Hand. Why Sign Came before Speech.* Wahsington, D.C.: Gallaudet University Press.

Tanner, J. E. and Byrne, R. W. (1996) Representation of action through iconic gesture in a captive lowland gorilla. *Current Anthropology*, 37/1: 162-73.

- (1999) The development of spontaneous gestural communication in a group of zoo-living lowland gorillas. In: *The Mentalities of Gorillas and Orangutans – Comparative perspectives*, ed. S. T. Parker, R. W. Mitchell and H. Lyn Miles. Cambridge University Press.

Tanner, J. E. (2004) Gestural phrases and gestural exchanges by a pair of zoo-living lowland gorillas. *Gesture*, 4/1: 1-24.

Terrace, H. S., Petitto, L. A., Sanders, R. J. and Bever, T. G. (1981) *Science* 211/4477: 87-88.

Thompson, R. K. R., Boysen, S. T. and Oden, D. L. (1997) Language-naive chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology*, 23/1: 31-43.

Tomasello, M. (1992) *First verbs: A Case Study of Early Grammatical Development.* Cambridge University Press.

- (1993) It's imitation, not mimesis. Commentary to Donald: *Origins of Mind. Behavioral and Brain Sciences* 14/4: 771-772.
- (1996) Do apes ape? In: *Social Learning in Animals: The Roots of Culture*, ed. C. Heyes and B. Galef. Oxford: Academic.
- (1999) *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard University Press.
- (2003) *Constructing a Language: A Asage-based Theory of Language Acquisition.* Cambridge, MA: Harvard University Press.

Tomasello, M. and Call, J. (1997): *Primate Cognition.* Oxford University Press.

Tomasello, M., Call, J. and Gluckman, A. (1997) Comprehension of novel communicative signs by apes and human children. *Child Development,* 68: 1067-1080.

Tomasello, M., Call, J. and Hare, B. (2003) Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends in Cognitive Sciences*, 7/4: 153-156.

Tomasello, M., Call, J., Nagell, K., Olguin, R. and Carpenter, M. (1994) The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates*, 35/2: 137-154.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (in press) Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences.*

Tomasello, M., Kruger, A. and Ratner, H. (1993) Cultural learning. *Behavioral and Brain Sciences*, 16, 495-552.

Tomasello M., Savage-Rumbaugh S. and Kruger, A. (1993) Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development*, 64: 1688-1795.

Tomasello, M. and Zuberbühler, K. (2002) Primate vocal and gestural communication. In: *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, ed. M. Bekoff, C. Allen and M. Burhgardt. MIT Press.

Trevarthen, C. (1992) An infant’s motives for speaking and thinking in the culture, In: *The Dialogical Alternative*, ed. H. Wold. Oslo: Scandinavian University Press.

Varela, F., Thompson, E. and Rosch, E. (1991) *The Embodied Mind. Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Veà, J. J. and Sabater-Pi, J. (1988) Spontaneous pointing behaviour in the wild pygny chimpanzee (*Pan paniscus*). *Folia Primatologica*, 69: 289-290.

von Hofsten, C. (1983). Catching skills in infancy. *Journal of Experimental Psychology: Human Perception and Preference*, 9: 75-85.

Whiten, A. (1998) Imitation of the sequential structure of actions by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 112: 270-81.

Whiten, A. and Byrne, R. W. (1988) Tactical deception in primates, *Behavioral and Brain Sciences,* 11: 233-244.

Whiten, A., Custance, D. M., Gomez, J.-C., Teixidor, P. and Bard, K. A. (1996) Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 110/1: 3-14.

Whiten, A., Horner, V., Litchfield, C. A. and Marshall-Pescini, S. (2004) How do apes ape? *Learning and Behaviour*, 32/1: 36-52.

Wilcox, S. (2004) Gesture and language: Cross-linguistic and historical data from signed languages. *Gesture*, 4/1: 43-73.

Wimmer, H., Hogrefe, G.-J. and Perner, J. (1988) Children’s understanding of informational access as a source of knowledge. *Child Development,* 59: 386-396.

Wittgenstein, L. (1953) *Philosophical Investigations*. Basil Blackwell.

Woll, B. and Kyle, J. (2004) Sign language. *Encylopedia of Language and Linguistics*. Oxford: Elsevier.

Woodruff, G. and Premack, D. (1979) Intentional communication in the chimpanzee: the development of deception. *Cognition,* 7: 333-362.

Vygotsky, L. S. (1962) *Thought and Language*. Cambridge MA: MIT Press.

- (1978) *Mind in Society. The Development of Higher Psychological Processes*. Harvard University Press.

Zlatev, J. (2002) Mimesis: The “missing link” between signals and symbols in phylogeny and ontogeny. In: *Mimesis, Sign and the Evolution of Language,* ed. A. Pajunen. University of Turku Press.

- (2003) Meaning = Life (+ Culture). An outline of a unified biocultural theory of meaning. *Evolution of Communication,* 4/2: 253-296.
- (2005a) What’s in a schema? Bodily mimesis and the grounding of language. In: *From Perception to Meaning: Image Schemas in Cognitive Linguistics,* ed. B. Hampe and J. Grady. Mouton.
- (2005b) The dependence of language on consciousness. Paper presented at the conference *Towards a Science of Consciousness 2005*, August 17-20, Copenhagen.

Zlatev, J., Persson, T. and Gärdenfors, P. (in press) Triadic bodily mimesis is the difference! Commentary to Tomasello et al. (in press).

Zuberbühler, K. (2000) Interspecies semantic communication in two forest primates. *Proceedings of the Royal Society of London. Series B Biological Sciences* 267: 713-718.