# Intersubjectivity, mimetic schemas and the emergence of language

Jordan Zlatev
Centre for Languages and Literature
Lund University
jordan.zlatev@ling.lu.se

**Abstract**

In this article I argue that intersubjectivity constitutes an essential characteristic of the human mind. First, I explicate what this means, contrasting intersubjectivity with the more common approach to social cognition based on the concept of "theory of mind". Furthermore, I propose a particular version of the thesis of the primacy of the shared mind based on the notion of *bodily mimesis:* a capacity to use our bodies in feeling the emotions of others, understanding their intentions, and eventually understanding and expressing communicative intentions. Bodily mimesis first takes place *between* people (and to a lesser degree, other higher social animals like apes and dolphins) but is gradually internalized as *mimetic schemas* (Zlatev 2005, 2007): preverbal concepts which possess a number of properties which can help explain the emergence of language as a conventional-normative semiotic system. In a dialectical fashion, intersubjectivity serves both as a prerequisite for its emergence, and is further developed by it, making human beings the quintessentially "intersubjective species". Finally, I suggest a partially novel approach to explaining autism.

Keywords: social cognition, imitation, bodily mimesis, representation, consciousness, autism

## 1. Introduction

"We are primarily *individuals*, living in our separate mental worlds, and occasionally connecting to other individuals through the imperfect means of language." From Descartes (1968 [1637]) and Leibniz (1973 [1787]) to Chomsky (1975) and Fodor (1983), this has been the predominant view of the human mind in the West. But this view comes with its costs: the spectre of solipsism, the gulf to "other minds", the difficulty to account for the possibility of mutual understanding, and even language itself. A reaction has been the retreat into "behaviour", of either the philosophical (Ryle 1949) or the psychological varieties (Skinner 1953), but such a stance throws out the baby – consciousness – along with the individualist bathwater (Maslin 2000; Baars 2003).

Not surprisingly, a number of thinkers from different traditions such as phenomenology (Husserl 1999 [1907], Merleau-Ponty 1945; Gallagher 2005), analytical philosophy

(Wittgenstein 1953; Davidson 2001; Hutto 2004), social-cultural psychology (Vygotsky 1978; Bruner 1990) and infant psychology and psychoanalysis (Stern 1985; Hobson 2004) have objected to *both* cognitive individualism and behaviourist anti-subjectivism. To the first – for distorting the reality of our lives as cultural beings and failing to account for the essentially social nature of language; to the second – for draining our lives of experiential content. Such thinkers have tried, in their various ways, to elaborate a third way between the two extremes, and this is the alternative of *intersubjectivity*. Basically, such an approach claims that we are fundamentally interconnected through shared emotions, practices, understandings, and (eventually) language – and on the background of this "shared mind" our individualities emerge like islands on a sea of primary interconnectedness.

In this article I clarify the intersubjective perspective on the human mind by first contrasting it with a more common approach to social cognition using the concept of "theory of mind" (e.g. Baron-Cohen 1995). I will then propose a particular version of the thesis of the primacy of the shared mind which views it as based on our capacity for *bodily mimesis:* a capacity to use our bodies as resonance boxes, so to speak, in feeling the emotions of others, understanding their intentions, and eventually for understanding and expressing communicative intentions. As in the famous Vygotskyan dictum of the interpersonal preceding the intrapersonal in development, bodily mimesis first takes place *between* people (and to a lesser degree, other higher social animals like apes and dolphins) but is gradually internalized as covert *mimetic schemas* (Zlatev 2005, 2007), which are used in thought as well as action – thought being a covert form of action.

My argument is that mimetic schemas constitute preverbal concepts which possess a number of properties which can help explain the emergence of language as a "socially shared symbolic system" (Nelson and Shaw 2002). The last feature discussed, their intersubjective nature, ties in with the first theme, and suggests a partially novel approach to explaining autism. I conclude by summing up the ways in which mimetic schemas help explain the ontogenetic emergence of language – which constitutes, in a dialectical fashion, both a product of intersubjectivity and a factor for the further development of *the shared mind* (Zlatev et al. in press).

## 2. Intersubjectivity vs. "theory of mind"

Intersubjectivity can be most generally defined as *the sharing and/or understanding of others' experiences*.[1] This includes not only beliefs and other "propositional attitudes", but phenomena such as emotions, attentional states and intentions. Crucially, it does *not* imply that mental states are initially *private*, and only secondarily, through inference or

[1] This includes pre-verbal as well as verbal experience ("shared meanings"), involving processes of affective, perceptual and reflective consciousness such as empathy, joint attention and folk psychology. Thus I include what others (and an anonymous reviewer) refer to as "socialness" within the concept of intersubjectivity. At the same time, it is possible to define *types* of intersubjectivity and to trace their development in phylogeny and ontogeny, as suggested in this essay and in previous work (Zlatev in press).

simulation, can be projected onto others. Rather, the assumption is that some basic forms of human consciousness are intersubjective from the start.

While this notion is well-established within the phenomenological tradition (e.g. Husserl 1999 [1907]; Merleau-Ponty 1962 [1945]; Gallagher 2005, Chapter 9), and in some of the developmental literature (e.g. Trevarthen 1979; Stern 1985), the dominant approach to social cognition within psychology and cognitive science is that of "theory of mind" (ToM). The difference between the two traditions is considerable. From the ToM perspective, the original privacy of mind is most often taken for granted and the question is how the child gains an understanding of "other minds". The two main alternatives are either through a conscious or unconscious *theory,* positing mental states and processes to explain observable behavior (e.g. Baron-Cohen 1995), or through – again conscious or unconscious – *simulation* of the other's mental activity within one's mind, and then "projecting" this onto the other (e.g. Goldman 1989). To account for ontogenetic development ToM models such as that of Baron-Cohen (1995) are forced to adopt innate "modules" such as "intention detection" (ID) and "eye-direction detection" (EDD) and building on these, at least a partially innate "shared attention mechanism" (SAM).

The existence of such structures is, however, questionable on philosophical, neuroscientific and evolutionary grounds (Hutto 2004, in press). The basic error of most ToM approaches (as typical for "computational" cognitive science, cf. Varela Thompson and Rosch 1991; Searle 1992) appears to be that they postulate mechanisms on the basis of introspective analysis of fairly advanced cognitive skills such as false-belief understanding – and then *project* these mechanisms onto the unconscious "sub-personal" level.

According to the perspective of intersubjectivity, the child's initial encounter with others is neither theoretical, consisting of beliefs and hypotheses about others' mental states, nor a matter of "simulating" these, but rather of perceiving these *directly* in engaged dyadic interactions, and furthermore that such primary intersubjectivity is our normal way of dealing with others even in adulthood:

> For we certainly believe ourselves to be directly acquainted with another person's joy in his laughter, with his sorrow and pain in tears, with his shame in blushing, with his entreaty in his outstretched hands … And with the tenor of his thoughts in the sound of his words. If anyone tells me that this is not "perception", for it cannot be so, in view of the fact that a perception is simply a "complex of physical sensations" … I would beg him to turn aside from such questionable theories and address himself to the phenomenological facts. (Scheler 1954, cited in Gallagher 2005: 228)

At the same time, such claims of "direct" or "primary" intersubjectivity need to be qualified in at least two ways. Firstly, the directness is phenomenological, i.e. on the level of *experience*, and not on the level of mechanism. There are certainly a number of complex neural structures and processes making intersubjective sharings possible, while at the same time segregating between "my own" from "your" actions. A number of plausible candidates for such mechanisms involving "mirror neurons" (Rizzolatti et al. 1996; Gallese, Keysers and Rizzolatti 2004) and "shared representations" (Georgieff and Jeannerod 1998; Decety

and Chaminade 2005) have been proposed during the past decade. Still, as long as the operation of such mechanisms is not experienced itself but only their outcomes, this does not invalidate the "phenomenological facts" referred to by Scheler in the quotation above. Furthermore, such structures and processes are *not* similar to the presumed ToM "modules" since they are both non-computational and cross-modal, rather than "encapsulated" (Fodor 1983).

Secondly, it is not the case that direct, interactional understanding of others *exhausts* our capacity to understand their minds. We are, after all, capable of *re-enacting* their actions, either overtly through imitation, or covertly in imagination and such processes are (phenomenologically) distinct from perception. In some cases we may even "theorize" in a detached way about their motivations and knowledge, once we have the semiotic resources given to us by a public language to represent these as "states" and "beliefs". At the same time, rather than simply contrasting such skills with (primary) intersubjectivity, it is possible to account for their development in terms of a progression of stages (e.g. Tomasello et al. 2005; Zlatev, Person and Gärdenfors 2005) and thereby to account for the evidence called upon by both simulation-theorists and theory-theorists without assuming innate modules.

Another difference between the perspective of intersubjectivity and that of ToM is that the first is closely related to the *body* (both the biological "living" and the phenomenal "lived" body) of the subject and of the Other. The classical phenomenologists argued that the direct perception of the mental phenomena, including the emotions, of others is made possible by one's "corporeal schema" which is pre-reflectively matched to the bodies of others, allowing seamless identification. The recent neuro-cognitive models mentioned earlier based on similar activations in premotor, prefrontal and parietal brain areas during performing, perceiving or imagining an action are compatible with the phenomenological analyses. However, such models are often formulated in terms of "neural simulations" (Gallese and Lakoff 2005), which is ambiguous between "subpersonal" unconscious processes and conscious imagination, and the two need to be distinguished. In the first case, the term "simulation" is superfluous, as argued by Gallagher (2005: 221):

> … since mirror neurons involve extremely good examples of intermodal perception (translating vision into proprioceptive and body schematic registrations), the most parsimonious simulationist account is that when we perceive another person's actions, that perception registers in the mirror system as already a first-person model of what such actions would be if they were the perceiver's own actions.

Apart from perception, action and imagination, mirror neuron systems are also implicated in the *imitation* of others' actions (Rizzolatti and Arbib 1998; Arbib 2003, 2005; Decety and Chaminade 2005). While phenomenologists have mostly tended to link intersubjectivity to bodily *perception*, other thinkers have tended to emphasize imitation as a more active mode of understanding others, as well as of oneself. Baldwin (1894: 42) famously asserted: "My sense of myself grows by imitation of you and my sense of yourself grows in terms of myself", proposing a model according to which imitation serves

as the primary means through which infants learn both about others, and develop self-consciousness through a dialectal process summarized by Zelazo and Lourenco (2003: 60) as follows:

> The dialectic starts with the presence of action that is (at least partially) outside of one's behavioural repertoire, and hence, viewed in terms of its outward or *projective* elements. By imitating, one comes to comprehend the *subjective* side of it; for example, one comes to appreciate the affect that accompanies it, or the effort involved. Once this happens, one tends automatically to eject this *subjectivity* back into the original behaviour.

Piaget (1951 [1945]) viewed imitation as the major source of deriving the first internal representations which he called "pre-concepts", suggesting the following developmental progression: (1) *sensorimotor imitation:* in which the model's action is imitated immediately following the stimulus; (2) *deferred imitation:* in which the performed action is a copy of an action removed in space and time, either of another, or of oneself; (3) *representational imitation:* in which "the interior image precedes the exterior gesture, which is thus a copy of an "internal model" that guarantees the connection between the real, but absent model, and the imitative reproduction of it." (Piaget 1951: 279). The pre-concepts that emerge through imitation can be said to constitute internalized, generalized imitations, which prior to being intra-personal have been inter-personal, in a way that is reminiscent of Vygotsky (1978).[2] Piaget was therefore not such as "individualist" as presented in most textbooks, and it is possible to argue that his notion of representation and symbolization rests on intersubjectivity (Zlatev 2005, see also Section 5.5).

Gopnik and Meltzoff (1993: 352), otherwise proponents of a non-innate, gradually constructed "theory-theory" of mind, argue that "normal children are innately endowed with the capacity to imitate others. This provides a social bridge between the newborn and caregivers." In earlier work (Zlatev 2000) I proposed that such an original "bridge" of pre-reflective intersubjectivity serves as the basis and necessary precondition not only for subjectivizing the Other, but of "objectifying" (i.e. explicitly conceptualizing) one's sense of self, moving from a largely unconscious *body schema* coordinating movements to a consciously accessible *body image* (cf. Gallagher 2005, for the distinction between the two concepts).

I

The term "imitation" is quite ambiguous in the literature, with some authors applying it to even the most simple acts of copying the motion of another, and others insisting on high-level cognitive processes such as the understanding of the model's goals/intentions. One way to resolve such disagreements is to regard imitation, like intersubjectivity, as a non-unitary phenomenon and as Piaget (1951) to see it on several different evolutionary and developmental levels. One of the most basic forms of imitation, *neonatal mirroring*

[2] At the same time, this typical sequence cannot be claimed to be a necessary invariant: sensorimotor imitation may be denied to children with motor disabilities, but they can nevertheless identify with others, and engage in covert imitation, i.e. imaginary re-enactment (see Section 5.4). Nevertheless, this would imply internalization (appropriation) of the inter-personal and would not prevent intersubjectivity and mimetic schemas from developing.

(Meltzoff and Moore 1978, 1983) does not require any explicit differentiation between the subject and the model. Thus, it corresponds to "primary" (Trevarthen 1979), or "proto-mimetic" (Zlatev, Persson and Gärdenfors 2005) intersubjectivity, which requires neither (true) representations nor simulations (cf. Gallagher 2005, Chapter 9). Importantly, such imitation has also been demonstrated in chimpanzee neonates (Myowa-Yamakoshi 2004) while "true" imitation, involving the copying of both the *means* and the *goal* of *novel* actions, is something in which apes may have some basic abilities (Whiten et al. 2004), but which is much more developed in our species (Donald 1991; Tomasello 1999; Arbib 2005). The relationship between imitation and intersubjectivity is a matter of current controversy: is understanding the "intentions" of others a prerequisite for imitation or vice versa (cf. Hurely and Chater 2005)? Using a broad, stage-based concept of imitation, allows us rather to view the two as co-evolving (Zlatev in press) and co-developing: "simple" forms of imitating the actions of others, overtly or covertly, allows for better understanding of their subjectivity and thus intentions. When children become experts in this, they also become better imitators, being able to distinguish between the means and goals of an action, and thus to engage in "true" imitation (Tomasello 1999).

It has been my proposal that the concept of *bodily mimesis*, constituting a generalization of imitation, imaginative re-enactment and gesture, can unify most of these perspectives and findings, and at the same time explain the possibility for language to arise (Zlatev 2000, 2002, 2003, 2005, 2007, in press). I will elaborate on this in Section 4, but prior to this we need to address the fundamental question "What is language?" and relate the answer to intersubjectivity.

## 3. The normativity of language and its bases

Language, often characterized as the only feature that distinguishes our species (e.g. Chistiansen and Kirby 2003), has been defined in widely different ways depending on the theoretical biases of linguists, who are proverbial for having quite contradictory conceptions of the nature of their object of study: a "hidden code" (Sapir 1928), a "mental organ" (Chomsky 1975), a "computational device" (Jackendoff 1983), an "instinct" (Pinker 1994) etc. What all such accounts miss is the essentially *social* nature of language, with *normativity* as a central characteristic (Itkonen 1978, 1997, 2003). As pointed out famously by Wittgenstein (1953) knowledge of language implies knowing certain *criteria of correctness* and such knowledge cannot be (radically) private, but must be shared within a community.[3]

Such sharing can take somewhat different forms, in a cline from being *mutually known* (Clark and Marshal 1983) to being *distributed* within the community (Putnam 1975). In the

[3] Cf. "… a person goes by [i.e. follows] a sign post only in so far as there exists a regular use of sign-posts, a custom… Is what we call "obeying a rule" something that it would be possible for only *one* man to do and to do only *once* in his life? … – To obey a rule, to make a report, to give an order, to play a game of chess are *customs* (uses, institutions)." (Wittgenstein 1953: §198, §199)

first case, all speakers of English have certain shared intuitions that, say: *An elephant is an animal* is a correct (truthful) assertion, while *An elephant is a number* is an incorrect (false) one. Since these intuitions are normative, i.e. intuitions about what people *aught to* say, speakers of English are justified in expecting them to be shared by other competent speakers of the language. In this way, the norms of English (or any other language) are not only known, but *known to be known*, i.e. mutually known, without the need for such knowledge to be "explicitly represented" in the minds of speakers. Thus, linguistic normativity suggests a resolution to the problem of the "psychological implausibility" of mutual knowledge (cf. Sperber and Wilson 1995), at least as far as semantics and grammar are concerned. In the second case there is a "division of linguistic labor": I know that I can for instance use the encyclopedia to find out what the exact differences between *oak* and *ash* may be. Somewhat intermediate are cases where I do have intuitions of correctness, but I am less certain, possibly due to dialectal variation, e.g. is it better to say "*on* my view" or "*in* my view"?

Despite such differences, in all these cases the inherently social, intersubjective and (in most cases) normative nature of language is irreducible to individual minds, and even less so to individual brains, as certain "physicalist" theories (e.g. Gallese and Lakoff 2005) persistently claim. After this (admittedly rather too brief) discussion, we may define language succinctly as *a conventional-normative semiotic system for communication and thought*. This may be an underrepresented definition in both linguistics and cognitive science, but is one that can be found reflected in the work of Lewis (1969), Itkonen (1978, 2003), Clark (1996) and Tomasello (1992, 1999).[4]

Given this characterization of language and that of intersubjecivity offered in the previous section, it becomes clear that there is an intimate relationship between the two. But how can we characterize this relationship more precisely? It is possible to argue, along with e.g. Bloom (2000) that without some basic skills of intersubjectivity, language would be unlearnable: "…it is impossible to explain how children learn the meanings of a word without understanding of certain non-linguistic mental capacities, including how children think about the minds of others." (Bloom 2000: 2). The problem with this statement and Bloom's theory is however that it employs the conceptual apparatus of "theory of mind", and often attributes unnecessarily complex understandings of others' mentality to pre-linguistic children. A multi-level, and multi-process account of intersubjectivity is immune to this, and is fully compatible with the claim that the process of learning a language itself, as a conventional-normative semiotic system, develops the social cognition of children to the extent of mastering the concept of (false) *belief* (Tomasello 1999; Hutto 2004). There is, for example, substantial evidence for a correlation between successful performance on language proficiency and false-belief tasks, with the first predicting the second rather than vice versa (Astington and Jenkins 1999; Lohmann and Tomasello 2003).

[4] Cf. "Linguistic symbols are social conventions that package cognition in a way that human beings have found useful for communication" (Tomasello 1992: 215).

In sum, the relationship between intersubjectivity and language seems to be one of co-development in ontogeny, and (possibly) co-evolution in phylogeny: a certain degree of intersubjectivity is necessary for making language possible, while the latter on its part contributes to the formation of the "mediated mind" (Nelson 1996). The crucial question is: what is it that makes language possible in the first place and starts the co-evolutionary and co-developmental snowball? This question is heatedly discussed in the literature on language origins nowadays (cf. Tomasello et al. 2005 and Commentary). In this essay I will address the question in the context of ontogeny, though see Zlatev (in press) for a corresponding perspective on evolution.

The first requirement for the *normal* acquisition of language is arguably a basic capacity for intersubjectivity involving identification with others, and a reciprocal emotional relationship with caregivers (Stern 1985; Hobson 2004). This does *not* mean that various degrees of linguistic proficiency cannot be obtained by high-functioning children with autism who lack such intersubjective engagements. It does, however, imply that their way of learning will follow other (and slower) developmental patterns than those of normal children, and that even those relatively few persons with autism who master language proficiently, would deviate in their use of language in social interaction. These predictions appear to be confirmed (Menyuk and Quill 1985; Bishop 1989; Tager-Flusberg 2000). On the other hand, if autism is linked directly to ToM, and a particular empirical (Bloom 2000) or philosophical (Lewis 1969) theory of meaning implies that language acquisition or use would be impossible without ToM, then high-functioning children with autism do present "counterexamples" to such theories (Glüer and Pagin 2003). However, if the relationship is rather between autism and intersubjectivity, both of which are *gradient* concepts, then what one would expect are correlations between the degree of the autistic syndrome and language impairment, which again appears to be the case (Tager-Flusberg 2000).

On a more cognitive level, it has been argued that in order to bootstrap into language, both lexically and grammatically, the child needs pre-linguistic *concepts,* rather than simply perceptual (or sensorimotor) schemas (Piaget 1951; Bloom 2000; Mandler 2004). In brief, the arguments are that language cannot be learned through ostensive pairings between expressions and actual events in the "hear and now" since (a) acts of ostension underdetermine the meaning of expressions, (b) referents are in most cases not perceptually given and (c) a large number of linguistic expressions do not refer to specific "basic-level" objects such chairs and dogs, but to actions (e.g. *run*), states (e.g. *all gone*), relations (e.g. *over*) or lack referential meaning entirely (e.g. *the*, *not*, *democracy*). Additionally, Fodor (1975) has famously argued that since language learning requires hypothesis formation and testing, these require on their part a representational medium, a "language of thought" in which to be formulated. Each one of these arguments has been extensively discussed and hotly contested (cf. Sinha 1988; Nelson 1996; Bermúdez 2005). Nevertheless, taken together they do pose a rather heavy burden for theories which aspire to link language and perceptual experience (Plunkett et al. 1992; Elman 1993) or conditions of use (Zlatev 1999) *directly*, i.e. without the mediation, in both language acquisition and use, of conceptual structure which is not (entirely) derived from language itself.

Mandler (2004) attempts to resolve this problem by proposing that children learn language on the basis of pre-linguistic *image schemas* such as ANIMATE MOTION, CONTAINMENT and PATH. The concept "image schema" derives from Cognitive Linguistics (Lakoff 1987; Johnson 1987) where it is claimed to play a pivotal role in the "grounding" of language. However, this concept and the semantic theories dependent on it suffer from a number of problems, mentioned below and elaborated on in Section 5. First, it is far from clear is an "image schema": different theoreticians attribute to it quite conflicting properties (cf. Hampe 2005; Zlatev 2005). Second, its leading exponents Mark Johnson and George Lakoff regard it as a non-representational – either interactional or neural – category, which leaves the representational (semiotic) character of language still to be explained. Third, these same authors regard it as part of the "cognitive unconscious" (Lakoff and Johnson 1999), while it can be argued that linguistic meaning is dependent on consciousness both ontologically and methodologically (Talmy 2000; Zlatev 2003, in press). Fourth, some of the hypothetical "image schemas", e.g. VERTICALITY, are quite abstract, in a way to make it difficult to see how they can be acquired independently of language. Fifth, while in principle multi-modal (or cross-modal), image schemas are implicitly or explicitly assumed to be based on *vision* (Mandler 2004), while blind children are not necessarily cognitively or linguistically impaired. Sixth, and most relevant for this essay: as most usually defined,[5] image schemas are assumed to be essentially individual sensorimotor categories, rather than intersubjective concepts.

In contradistinction, I have recently proposed (Zlatev 2005, 2007) the notion of *mimetic schemas*, which are fairly specific, cross-modal, consciously accessible representations based on imitation, and largely shared within a (sub)culture. In the next section, I will provide a definition and in the following one, will explicate each one of their defining properties, which – I suggest – render bodily mimesis and mimetic schemas eligible as a basis for the emergence of language and the development of the shared mind.

## 4. Bodily mimesis

In his influential synthetisizing monograph on human origins, Donald (1991) argues that a form of cognition crucially based on *mimesis*, and a corresponding "mimetic culture" characterizing the primitive societies of *Homo ergaster/erectus* mediated between the "episodic" cognition of the common ape-human ancestor and the evolution of language as a dominant mode of human communication. Mimetic representations are according to Donald "conscious, self-initiated, representational acts that are intentional but not linguistic" (ibid: 168). More specifically, mimesis involves a number of different skills such as (true) imitation, the re-enactment of actions in imagination – allowing for explicit memory, planning and rehearsal, and the use of iconic and indexical gestures for intentional

[5] Cf. "An image schema is a recurring dynamic patterns of our perceptual interactions and motor programs that gives rise to coherence and structure to our experience" (Johnson 1987: xiv), "The most useful way of understanding image schemas is to see them as mental representations of *fundamental units of sensory experience*" (Grady 2005: 44).

communication. Others have argued for a similar "mimetic stage" in ontogeny, but have proposed quite different interpretations (Nelson 1996; Zlatev 2002, 2003), making it clear that the concept of mimesis requites a more precise definition. Building on Donald's work, but taking into account some of the recent evidence on social neuroscience mentioned in the previous section, as well as evidence on the mimetic capacities of non-human primates, summarized by Zlatev, Persson and Gärdenfors (2005), the concept of *bodily mimesis* can be defined as follows:

> **Def:** A particular bodily act of cognition or communication is an act of **bodily mimesis** if and only if:
>
> a) It involves a cross-modal mapping between *exteroception* (i.e. perception of the environment, normally dominated by vision) and *proprioception* (perception of one's own body, normally through kinesthetic sense);
> b) It is under conscious control and corresponds to – either iconically or indexically – to some action, object or event, while at the same time being *differentiated* from it by the subject;
> c) The subject *intends* the act *to stand for* some action, object or event for an addressee (and for the addressee to recognize this intention);
> d) Without the act being conventional-normative, and
> e) Without the act dividing (semi)compositionally into meaningful sub-acts that systematically relate to each other and other similar acts.

This definition clarifies a number of issues on the relationship between bodily mimesis and related but distinct phenomena. Condition (a) states that an exteroception-propioception mapping (possibly supported by the "mirror neuron" system) is a necessary condition for bodily mimesis. However, it is not sufficient. If only condition (a) is fulfilled, as is the case in contagion and neonatal mirroring, the subject will be capable only of *proto-mimesis*, but not bodily mimesis proper.

Condition (b) states that a mimetic act needs to be *volitional* and *representational*, as in Donald's original definition, and explicates the notion of "representation" in line with Piaget's (1945) criterion of differentiation between "signifier" and "signified" from the subject's point of view (cf. Sonesson 2007), adding the requirement that the signifier is a bodily act. Piaget's example of an infant opening and closing her mouth to model the opening and closing of a matchbox would be an example of an *iconic* correspondence. Children's acts of pointing for themselves in order to help guide their attention (Bates, Camaioni and Volterra 1975) would qualify as *indexical* mimetic acts.

However, if condition (c) is not fulfilled, the acts would involve *dyadic* mimesis. Condition (c) introduces the necessary *triadic* element in order to make bodily mimesis

communicative: the representation or sign[6] is intended to be recognized as such by an addressee, along with the communicative intention itself. This introduces a Gricean element of intentional communication (Grice 1957), without minimizing the semiotic (representational) aspect of meaning as in Relevance Theory (Sperber and Wilson 1995). An example of an iconic sign that fulfills all three conditions is the *miming* of eating by pretending to move a spoon to one's mouth (made behind a glass door) in order to communicate to a colleague that it is time for lunch. An indexical mimetic sign would be, for example, a paradigmatic form of *declarative pointing* (Brinck 2003).

Finally, if an act fulfills not only (a)-(c), but also the positive versions of (d) and (e) – with conventionality-normativity and (semi)compositionality – it is not mimetic, but rather *symbolic,* since by adding these properties the border between the mimetic and the "post-mimetic" has been traversed, as is the case in language and possibly certain rituals. Notice that conventionality does *not* imply *arbitrariness*: a large proportion of the signs of signed language such as ASL, and some of those in spoken languages have iconic or otherwise *motivated* meanings, without this counteracting either their conventional or normative character (cf. Zlatev 2003). The transition between (triadic) mimesis and signed language can be observed either in macro-settings such as the emergence of Nicaraguan Sign Language over the span of three generations of signers (Senghas, Kita and Özyürek 2004) or in micro-settings, where a signer, in cohort with his addressee, has been observed to transform a mimetic gesture into a "micro-conventional" sign over the course of a single narrative (Taub 2005).

A few final clarifications need to be made. Both bodily mimesis and language are realized through specific acts, which are originally *overt*, but both can be *internalized* and used for thought (Vygotsky 1962, 1978; Piaget 1945; Donald 1991). But internalized, or *covert mimesis*, corresponds to what was discussed in the previous section as the "mental simulation" of actions, or "action images" (Jeannerod 1997). What is essential, however, is to distinguish this form of *conscious* imagination from the currently popular notion of "simulation" which conflates consciously accessible from "subpersonal" processes (e.g. Barsalou 1999; Gallese and Lakoff 2005). According to the definition of bodily mimesis, such unconscious processes would satisfy condition (a), and would therefore, be proto-mimetic. Also to emphasize once again, the way the term *representation* is used in the definition of bodily mimesis, it presupposes a differentiation between expression and content, or mental image and (perceptual) reality *from the standpoint of the subject* (Piaget 1945; Sonesson 2001, 2007), implying conscious awareness of the representation as such.

## 5. The properties of mimetic schemas

On the basis of the concept of bodily mimesis, the second central notion of the theory proposed in this article can be defined as follows:

[6] Similar to e.g. Sonesson (2007), I regard the terms *representation* and *sign* as synonymous (once properly defined). Since these are arguably the pivotal terms of cognitive science and semiotics, respectively, and identifying the two should contribute to the unification of these two different traditions.

**Mimetic schemas (Def):** *Categories of acts of overt or covert bodily mimesis.*

Prime examples of mimetic schemas with more or less universal status are those which categorize "unbounded" *activities* such as CRYING, RUNNING, CRAWLING and FLYING and "bounded" *actions* such as GRASP-X, PUSH-X, KICK-X and PUNCH-X. Other mimetic schemas will be more or less *culturally specific* such as ICE-SKATING and SHOOT-X.

My contention is that mimetic schemas constitute pre-linguistic concepts which are (a) representational, (b) accessible to consciousness, (c) intermediately abstract, (d) proprioceptively based and (e) pre-reflectively shared. I will explicate these five properties, contrasting mimetic schemas with the construct of "image schemas" when appropriate.

### 5.1 Representational

Mimetic schemas are *overt or covert representational structures*, but to make this statement meaningful, we must first define the concept of "representation" which has been overextended in much of cognitive science (cf. Newton 2003; Sonesson 2007). A representation is here understood as a structure that consists of three parts: an *expression* that stands for a given *content* for a given *subject*. Thus defined, it is identical with the classical definition of a *sign.* A clear example of a representation is a picture: the depicted apple cannot be eaten, but it represents (in this case *iconically*) an apple that can. The painting itself is the expression, and it is different from, at the same time as it corresponds to, something else. Whether this "something else" is a real specific apple, a generic apple, or an imagined apple is less essential than the *expression-content structure* itself. What "connects" the expression and the content is a process of *interpretation*: representations do not exist by themselves, but only for *someone* (cf. Sinha 1988, Chapter 2).

Pictures as well as iconic gestures are *overt* representations, but internalized or *covert* mimetic schemas correspond to what has traditionally been called "mental representations". This is the most controversial notion. "Classical", first-generation cognitive science postulated an overflow of (unconscious) mental representations: in thought, in language, in perception, in practical action (Gardner 1985). Some representatives of second generation, "embodied" cognitive science (Varela, Thompson and Rosch 1991; Johnson and Rohrer 2007) take the opposite view and more or less reject the notion of representation altogether. From the present perspective, these extremes are equally misguided: to say that a subject has a mental representation is to say that he or she can (a) *differentiate* between the expression and content, and (b) see the first as *corresponding* to the second, as in the picture example mentioned above, except that the mimetic schema is perceived in reflective rather than perceptual consciousness. One may ask: Who is the "someone" doing the differentiation and finding the correspondence? There are three types of answers to this question:

(a) An unconscious "subpersonal" (neural) mechanism. The problem with this answer is that postulating such a "homunculus" leads to infinite regress: we need to account

for the ability of the homunculus to "see" the expression and content, and "figure out" that the first stands for the second, possibly engaging in a mis-representation. But then we need to account for the mental representations in *its* head, *ad infinitum* (cf. Edelman 1992).

(b) The standard reply to the above objection within (computational) cognitive science is that the digital computer (or its central processor) operates on "internal representations" and this constitutes and existence proof of the invalidity of the homunculus argument. This reasoning however runs into other problems since such "representations" or "symbols" are meaningful only for someone *else* than the system that is actually using the symbols. But then the representation is not *intrinsic* to the system but to the programmer, or whoever else is doing the "interpreting" (cf. Searle 1992, 1999).

(c) The *subject* himself or herself, i.e. the conscious person who *experiences* the mental representation as such, i.e. distinguishes it from perception, and imbues it with a qualitative "subjective" tone.

To understand this last, and I believe only truly valid, sense of representation I ask you to close your eyes and *imagine* an apple: would you confuse your imagined apple with the one it represents? On the other hand, when you see an apple, do you think of your perception as a "representation" of an apple – or as the apple itself? The answers to these questions should be obvious. It is these basic facts about our experience that have been surprisingly overlooked by so many, which lead us to the conclusion that there are no mental representations in perception, but only in imagination (Piaget 1945; Gibson 1979; Sonesson 2007; Zlatev 2007). Again, this conclusion is based on phenomenological facts and does not state anything about the "underlying" neurological processes, whose nature is an empirical matter.

This account of mimetic schemas as (mental) representational structures stands in stark contrast to the anti-representationalism of (some) proponents of image schemas.

> As we said in *Philosophy in the Flesh*, the only workable theory of representations is one in which a representation is a flexible pattern of organism-environment interactions, and not some inner mental entity that somehow gets hooked up with parts of the external world by a strange relation called 'reference'. (Johnson and Lakoff 2002: 249-250)

> Image schemas are thus part of our non-representational coupling with our world, just as barn owls and squirrel monkeys have image schemas that define their types of sensorimotor experience. (Johnson and Rohrer 2007)

There is a serious problem for such theorists to the extent that they purport to provide an evolutionary and ontogenetic basis for language. Sensorimotor interaction is an inherently non-representational notion, while language *is* representational in two different, though related, respects: (a) it has the expression-content structure of *signs* and (b) statements are *about* (real or imagined) states of affairs (Searle 1999; Zlatev 2007). If image schemas such as PATH and CONTAINER, postulated to constitute the meanings of spatial prepositions such as *from* and *in*, are of the same kind as those of "barn owls and squirrel monkeys" then the representational character of language itself would be lost. Alternatively, one would need to defend the claims that representational cognition is only made possible by

language. However, there is strong evidence against this in both developmental psychology (e.g. Mandler 2004) and evolutionary theory (e.g. Donald 1991; 2001): neither pre-linguistic infants nor apes are trapped in the "here and now" but can engage in differed imitation and planning for future actions.

Mimetic schemas such as GRASP-X, JUMPING etc, which can be mentally rehearsed and combined, constitute a form of non-linguistic representational thought that can plausibly be assumed to play the role of a "bridge" between pre-representational sensorimotor cognition and language in both ontogeny and phylogeny (Donald 1991, 2001; Zlatev 2002, 2005; Arbib 2003, 2005). While there appear to be other forms of mental representations than mimetic schemas, such as memories of specific *episodes*, mimetic schemas are more abstract than so, and have a combinatory potential that approaches that of language, and therefore constitute a plausible "exaptation" for the evolution of language, thereby providing the proverbial "missing link".

### 5.2 Accessible to consciousness

As pointed out already, mimetic schemas require an irreducible reference to consciousness for establishing their representational character, as well as for their internalization. While it is still the case in many quarters that the notion of consciousness is regarded with trepidation, this is more of a hangover from behaviourism than a rationally justified reaction. Over the 20 years or so, consciousnesses has become both philosophically and scientifically respectable and there has been clear progress in both conceptual issues, such as elucidating the "hard problem" of the irreducibility of qualitative experience (Chalmers 1997) and methodological issues of combing "first-person" and "third-person" data in studying cognition and consciousness (Lutz and Thompson 2003). Such studies have clarified distinctions between different modes or kinds of consciousness, e.g. *affective, perceptual* and *reflective* consciousness, and have contributed to the understanding the neural underpinnings of experience (Edelman 1992; Georgieff and Jeannerod 1998; Damasio 2000).

Mimetic schemas involve reflective consciousness, since they can be *accessed* and *thought about* independently of whatever they represent, but they also contain affective aspects since they have what phenomenologists refer to as *affective tone* (Thompson 2001). To see this last point is sufficient to appreciate the representational and affective differences between the schemas KISS-X and PUNCH-X. As structures of reflective consciousness mimetic schemas constitute part of our *declarative* knowledge, as opposed to *procedural*, sensorimotor skills such as bicycle-riding.[7] This distinction is emphasized by Mandler

[7] It should be noted that by referring to mimetic schemas as "structures" I am in no way implying that they need to be *static*, and in this way am not presupposing a structure/process duality (cf. Zlatev 2005). In this sense my construct is similar to Newton's (1996, 2003) *basic action schemas*, which are also claimed to be both representational and (potentially) conscious. A difference, however, is that mimetic schemas can be said to represent categories, i.e. types of actions, which Newton views as "unnecessarily abstract and metaphysically confusing" and suggests that only "actions schemas *in use* can be said to represent the action that they facilitate" (Newton 2003: 188, original emphasis). However, such parsimony would prevent us from

(2004), who consistently distinguishes between the two sorts of knowledge, and corresponding forms of learning:

> Procedural knowledge, both perceptual and motor, is inaccessible to consciousness. … In spite of taking in lots of information at once … it is also relatively slow to learn, and learning is accomplished by associative strengthening, typically over a number of trials, as in operand conditioning or perceptual schema formation. It aggregates frequency information. […] Declarative or conceptual knowledge, in contrast, is accessible to awareness and is either describable in language, or, with a little analytic training, by drawing. It requires attention to be encoded into this format; this means that it is selective. (Mandler 2004: 55)

Declarative knowledge is both representational and consciously assessable and as such can be properly regarded as *conceptual.* In this sense mimetic schemas are pre-verbal concepts, while sensorimotor (or other purely interactional) schemas are pre-conceptual. But even the latter are hardly explicable as structures of the "Cognitive Unconscious": "…the realm of thought that is completely and irrevocably inaccessible to direct conscious introspection" (Johnson and Lakoff 1999: 12). In a recent publication, Johnson (2005) admits that by regarding image schemas as (unconscious) structures "you lose, or at least overlook, the very thing that gives image schemas their life, motivating force, and relevance to human meaning, namely, their embeddedness within affect-laden and value-laden experience" (ibid: 27). But this is nothing else but an aspect of affective consciousness, and thus an aspect of "phenomenological embodiment" (Lakoff and Johnson 1999), rather than the Cognitive Unconscious. Johnson's (2005) problem of how to include (affective) consciousness into image schematic structure, does not appear with respect to mimetic schemas, since mimetic schemas are defined as "categories of acts of… bodily mimesis", e.g. as generalizations from specific bodily actions, and each schema is characterized by a different type of emotional-proprioceptive tone (Thompson 2001).

To summarize, mimetic schemas require consciousness for (a) determining their representational character, i.e. they stand for something for a conscious subject; (b) they can be thought about (in the absence of language) as structures of imagination, and thus constitute a basic form of reflective consciousness and declarative knowledge and (c) they carry with them the affective components of real-life actions, and are thus endowed with affective tone.

### 5.3 Intermediately abstract

As pointed out earlier, mimetic schemas possess a level of abstractness that is greater than that of specific *episodes* (Donald 1991) or of specific *action schemas* (Newton 1996), but more specific than the one usually attributed to image schemas. The "proper level" of abstractness of image schemas in hotly contested in the Cognitive Linguistics literature. In addressing the issue of (image) schema-abstractness Grady (2005) distinguishes between three hierarchically ordered levels:

---

allowing mimetic schemas to play the role of *semantic representations* that can be combined as well as shared between different members of a community (see Section 4.3 and 4.5).

1. Specific schemas such as UP, for which Grady reserves the designation "image schema".
2. More abstract schemas such as MORE, which Grady calls "response schemas" since he regards them as being the outcome of "primary metaphors", defined as mappings from concrete and specific domains to abstract and fuzzy one due to correlations in experience in early childhood.
3. "Super-schemas" such as SCALAR PROPERTY, which capture the shared structure between (1) and (2)

Mimetic schemas such as CLIMBING and JUMPING are (at least) one level below Grady's first level. Being so close to perceptuomotor experience, it is fairly straightforward to account for their emergence through a process of generalization from instances of the corresponding acts – given normal generalization capabilities. However, inadequate generalization of perceptual experience is one of the features (and explanations) of autism (Plaisted 2001) and hence mimetic schemas may indeed be difficult to acquire for children with autism (cf. 4.6 below).

It is *much* more difficult to see how more abstract structures such as UP could arise simply by abstracting from the specifics of particular actions or events, as distinct as balloons rising, monkeys climbing, rockets lifting, gazes shifting their angle etc. The learning task would be simplified if *language* plays a key role in the generalization process, for example, if the child is exposed to the expressions *up* or *down* in conjunction with actions involving motions in the corresponding direction. It is characteristic that in Regier's structured connectionist model of the learning of spatial concepts (Regier 1996), it was necessary to provide such linguistic "labels" and furthermore to learn several different terms in parallel, e.g. *in*, *out* and *through*, which thereby provided "implicit negative evidence" for one another; otherwise the model failed to converge. As well-known, there is substantial cross-linguistic variation in the spatial semantic systems of different languages and depending on the "linguistic input" children form different spatial concepts from the onset of language acquisition (Bowerman 1996). For example, it has been shown that while both English and Korean infants are sensitive to the contrast between actions in which there is TIGHT-FIT vs. LOOSE-fit between the object of the action and the object in or on which it is placed, only Korean adults make this distinction in various categorization experiments. This can be explained by the fact that Korean has verbs which code this distinction: "putting in/on/together that result in a fitting relationship (KITTA) from those that result in loose containment (NEHTA) or surface contact (NOHTA, PWUTHIATA)" (Choi and Bowerman 1991: 92). This argumentation implies that image schemas even of the most concrete type in Grady's hierarchy cannot be the sole "ground" for language, since they are – at least in part – learned through it. Mimetic schemas such as RUNNING and GRASP-X, on the other hand, can play such a role since they are (plausibly) acquired pre-verbally.

At the same time as being "close" to actual bodily experience, mimetic schemas possess a considerable amount of internal structure which can facilitate language acquisition. One thing that has so far been implicit, reflected in the notation used, is *aspectual structure*: mimetic schemas expressed as VERB–ING are representations of unbounded *activities*,

while those given in the formula VERB–X are representations of goal-directed bounded *actions*. Aspect, whether lexically (i.e. "Aktionsart") or grammatically encoded, is a universal property of language and while languages differ substantially in their "tense and aspect systems" (cf. Dahl 1985) the distinction between activities and actions is always observed. Durst-Andersen (1992) argues that in their basic forms, verbs denote only three kinds of entities: states (e.g. *see*), activities (e.g. *eat*) and actions (e.g. *hit*), and interestingly the first two appear fairly early in child language acquisition, and then in present tenses, while action verbs come rather later, typically in past tense (ibid: 21). Durst-Andersen uses such observations to argue that while states and activities can be perceived directly, actions are "mental constructs", consisting two sub-situations: an activity and a state, connected be the relation of *telicity* which is not perceived but rather *conceived*: "we find state situations in reality (e.g. a person sitting on a chair) and activity situations in reality (e.g. a person jumping), but we find no (genuine) actions situations in reality…" (ibid: 61). This analysis is applied successfully to Russian aspect, and the distinction between states/activities and actions is convincing. Still, it seems that Durst-Anderson is operating with a rather narrow and static notion of "reality", which (without getting into metaphysics) hardly corresponds to what is "real", meaningful and salient to the pre-linguistic child. As argued earlier in this essay, due to our capacities for intersubjectivity, the goal-directed actions of others will (normally) be perceived directly and non-inferentially, no less so than their states and activities, once the child has formed corresponding action and mimetic schemas. While mimetic schemas for activities like RUNNING and actions like HIT-X are not identical with the corresponding verb meanings, which with their conventional-normative character are post-mimetic entities (cf. Section 3), they are rather straightforward candidates to ground the latter in ontogeny, and along with this the corresponding semantic distinction concerning aspect.[8]

In being more specific than the standard examples of image schemas such as SOURCE-PATH-GOAL, mimetic schemas are predominantly *iconic* representations, while the latter are often explicitly claimed *not* to be iconic (Mandler 2004, 2005). Iconic mimetic schemas *resemble* the represented activity or action in terms of features such as SPEED and DIRECTION as witnessed in *iconic gestures*. The great advantage of iconic, as opposed to arbitrary representations, is in the *transparency* of their content. Mimetic schemas thus do not need extra "rules of interpretation" in the manner of symbols (cf. Putnam 1975). When applied in a particular context – either in thought, or in communication through miming and gestures – there will naturally be a process of interpretation by both the gesturer and audience. Still, this process can be seen as one of *re-enaction*, rather than a convention that needs to be known in advance.

[8] The relative delay in the emergence of action verbs compared to activity verbs could be explained by their relative referential ambiguity, rather than their greater cognitive complexity. As Durst-Andersen (1992: 62) remarks: "A action manifests itself either as an **event**, i.e. as a state situation conceived to be caused by a certain prior activity situation [e.g. *He was killed.*], or as a **process** [e.g. *He gave her a flower (so that she would be glad)*.], i.e. as an activity situation conceived to be intended to cause a future state".

There is also at least one mimetic schema which is *indexical* rather than iconic. As noted in Section 3, an act of bodily mimesis may correspond to an action, object or event, at the same time as it is differentiated from it, on the basis of indexicality, i.e. spatiotemporal contiguity or factorality (Sonesson 1989; 2007). Reaching and "imperative" pointing which develops from it by ontogenetic ritualization (Tomasello 1999; Brinck 2003) involve indexicality and are attested in many non-human primates (Leavens and Hopkins 1999) and 9 month infants. However, since these are basically procedural skills aimed at achieving a practical goal, either without or with a bystander, there is no real differentiation between the act of pointing and its *meaning*. Therefore they are not true indexical signs, i.e. representations. On the other hand, "declarative", or rather *referential* pointing, in which the meaning of the gesture is not to solicit a desired object, but to focus the attention of the addressee to some aspect of the environment for communicative purposes, is a sign/representation in this sense (Zlatev, Persson and Gärdenfors 2005). Once the child acquires the ability to engage in referential pointing, through imitation rather than ritualization (Bruner 1983; Tomasello 1999), we may say that the child has acquired the mimetic schema POINT-TO-X, with its intersubjectively shared meaning. Evidence for this is that around their first birthday children both point referentially (e.g. at airplanes or other distal objects) and understand the pointing of others. Such pointing is a human universal, though the actual shape of the gesture – with the whole hand, the index finger, the head and mouth, or a combination of these – is conventionalized in different ways in different cultures (cf. Kita 2003). Furthermore, by combining referential pointing and an iconic gesture, or more abstractly POINT-TO-X + ICONIC SCHEMA-X, one has the minimal building blocks for expressing a *predication*. All this structure can exist pre-linguistically and can serve as the ground for the development of language in childhood.

### 5.4 Proprioceptively based

Image schemas, like mimetic schemas, are usually understood as *cross-modal* (as opposed to amodal) structures involving sensory-motor coordination. This formulation, however, avoids the question which sensory modalities are most *essential* in their formation. Is it possible, for example, to "subtract" two or even three sensory modalities, and expect that the remaining ones will be sufficient for the formation of such cross-modal schemas? Is there any one that is indispensable? The modality that is most often assumed to be most basic, either implicitly or explicitly so, is *vision.*[9] This is in itself not implausible given the relatively dominant role played by vision in the primate and even more so human brain (Watt 1991). However, some studies of *congenital blindness* show only slight delays in cognitive development and language onset, but otherwise a completely normal developmental pattern in language acquisition (Landau and Gleitman 1985). Other studies do show that congenital blindness can be a contributing factor to symptoms that resemble autism (Hobson, Lee and Brown 1999). But even so, it appears quite clear that blindness *predisposes for* rather than directly causes the abnormal developmental patterns. It is noteworthy that the blind child studied by Landau and Gleitman received a high degree of

[9] Mandler (2004, 2005), for example, argues that "the spatial information most crucial to human concept formation is delivered primarily by the visual system (ibid: 149).

haptic stimulation and linguistic interaction from her caregivers, which seemed to compensate for her lack visual input. While many questions on the underlying causality of the developmental processes involved in such clinical evidence remain unanswered, a rather straightforward conclusion is that while vision is undoubtedly a very important source of experience for normal children, it cannot be a *necessary* ground for either conceptual development or language.

Another hypothesis, deriving from the emphasis on sensorimotor activity in Piaget's developmental theory (Piaget 1954), as well as Johnson's (1987, 2005) and Grady's (2005) description of image schemas as emerging, above all, from concrete physical experience, is that actual physical (kinesthetic) experience is a necessary condition for cognitive (and eventually linguistic) development. This may also seem to be a predication made by the theory presented in this article, if assumed that overt bodily mimesis (in imitation and gesture) must always precede the formation of covert mimetic schemas. We can thus formulate the following prediction: if image or mimetic schemas are acquired through *actual* physical experience, and they are a necessary prerequisite for the development of (spatial) cognition and language, then severely motorically impaired children, such as those suffering from *spinal muscular atrophy* (SMA) would be expected to be at least cognitively and linguistically retarded. However, it has been shown that children with SMA do not differ in their performance on tasks involving spatial cognition compared to healthy controls (Riviére and Lècuyer 2002, 2003) and display normal IQ and language skills. Therefore it follows that either image/mimetic schemas are not acquired through actual physical experience or that they are not necessary for cognitive and linguistic development.[10]

The definition of bodily mimesis presented in Section 4 offers a way to resolve this difficulty. Since mimetic schemas are "categories of acts of … bodily mimesis" it follows form point (a) of the definition that they are based on a cross-modal mapping between exteroception (dominated by vision, unless the child is blind) and *proprioception* (dominated by kinesthetic sense), a mapping that is hypothetically mediated through a pre-motor-parietal-temporal system for action-proprioception-perception mapping (Iacoboni 2005). If a new-born child is deprived from either proprioception or from the proprioception-perception mapping, the prediction is that it will be quite helpless both motorically and cognitively, and would hardly even survive. However, if proprioception is (largely) lost in mature life, the effects should not be as negative. Gallagher (2005) describes extensively a well-known patient, IW, who has lost tactile and proprioceptive input from the neck down as an adult but with much effort has learned to consciously control his movements, especially through visual guidance. In Gallagher's (2005: 37-38) analysis this patient performs this through the *body image*, which is defined as "a (sometimes conscious) system of perceptions, attitudes, beliefs and dispositions pertaining to one's own body." What Gallagher suggests is impaired in the patient is the *body schema*: "a system of sensory-motor processes that constantly regulate posture and

[10] Mandler's (largely) vision-based account is fully compatible with these results, since it downplays the role of tactile and kinesthetic experience. But then it is problematic due to the results from congenital blindness.

movement – processes that function without reflective awareness or the necessity of perceptual monitoring" (Gallagher 2005: 37-38) which depends crucially on proprioception. The two systems normally interact, in development as well as in motoric performance, but can be doubly disassociated as shown from studies of hemineglect, where the patients appear to have an intact body schema, but are more or less "blind" to part of their bodies and visual field phenomenologically, even though they may still react to stimuli in that sub-field on the basis of non-conscious processes.

Given these concepts, we can hypothesise that bodily mimesis *originates* by linking the body schema and the external environment and above all other people, but since the body schema is not under conscious control, on this level we can only speak of *proto*-mimesis. Full bodily mimesis and mimetic schemas, characterized in this section as *conscious, dynamic representations*, must rather involve the body *image.* This serves as a kind of *virtual* body and appears to correspond to the "mimetic controller" hypothesized by Donald (1991). Thus IW would not be impaired from using mimetic schemas, as here defined.

Returning to the children with SMA, while they are motorically impaired, this does not need to involve the "innate" (i.e. present at birth) aspects of their body schemas, i.e. the ones responsible for neonatal imitation and for identifying with con-specifics (Gopnik and Meltzoff 1993; Gallagher 2005). This may be sufficient to bootstrap the development of their body images, allowing them to form mimetic schemas of the *covert* type. In other words, they would have been able to perform mental "simulations" of actions and events, matching these to the bodily motions of others, through a "shared representations" system, internalizing not their own actions, but those of others. This has implications for the basically intersubjective character of mimetic schemas, explicated below.

### 5.5 Pre-reflectively shared

The final characteristic of mimetic schemas brings us back to where we began: intersubjectivity. Image schemas, as nearly most constructs proposed within cognitive science, are usually conceived of as individual phenomena (cf. Sinha 1999), even when their "interactional" as opposed to subjective character is emphasized (Johnson 1987; Lakoff 1987; Lakoff and Gallese 2005). But as argued in Section 3, language is essentially a social, conventional-normative phenomenon (cf. Itkonen 2003). So, on the one hand we have a conceptual problem: how do we unite individual, "embodied" cognition and super-individual semantic and grammatical norms? At the same time we have an empirical problem: 18-month-old children who begin using language productively lack the meta-linguistic capacity for establishing full-fledged conventions; how do children therefore move from the sensorimotor to the symbolic (i.e. conventional-normative and systematic) level to learn a language as a "socially shared symbolic system" (Nelson and Shaw 2002)? Part of the answer may lie in the fact that language use itself teaches children the concepts of conventionality and normativity, through the negotiations of discourse (Tomasello 1999) and through the narratives that children are exposed to (Hutto 2003; in press). But this can not be the whole story, unless one wishes to claim that children's entrance into language is characterized by simple (behaviourist) "usage" rather than intersubjective *meaning* – and

this is implausible given the facts of pre-linguistic referential pointing and early language acquisition (Sinha 1988; Baldwin 1995; Bloom 2000).

The attempt to resolve these problems was the initial motivation that brought me to the concepts of bodily mimesis and mimetic schemas (Zlatev 2002, Zlatev 2003). My contention is that they can help to provide the necessary link between individual embodied cognition and collective representations, i.e. culture, due to the following considerations. Mimetic schemas will be overwhelmingly shared among the members of a community who engage in close face-to-face, or rather body-to-body interaction. A particularly important form of such interaction consists of (overt or covert) *imitation.* This yields what Arbib (2003, 2005) calls *representational parity* and the content of a mimetic schema, e.g. JUMPING, will be similar for the one who performs the act and for the one who observes it, imitates it, and internalizes it. As pointed out in Section 2, imitation was hypothesized by Piaget (1951 [1945]: 280) to serve as the route to developing the first (true) representations in children: "Imitation, with the help of images, provides the essential system of "signifiers" for the purpose of individual or egocentric representation". The latter term requires clarification; Piaget uses the term "egocentrism" to refer to a "failure to differentiate between the ego and the group, or confusion of the individual view-point and that of others" (ibid: 290). In other words, it is more appropriate to say that mimetic schemas are *pre-reflectively shared*, i.e. intersubjective, rather than "individual". In sum, since the child's mimetic schemas derive from imitating – overtly or covertly – salient activities and actions in the cultural community, both their representational and experiential content will be "shared" with those of his caregivers and peers.

Furthermore, since mimetic schemas are "categories of acts", arrived through by generalization, they will be relatively stable and delineated. There will thus be a limited set of mimetic schemas within a specific cultural community, and their *parity* can be not only on the level of individual actions, but of *types*.[11] As pointed out in Section 4.3, mimetic schemas possess a considerable amount of semantic structure. In both of these respects – sharing and structure – mimetic schemas can serve as a bridge to developing true *symbols*, which are (a) not just shared but *known to be shared* (i.e. conventional) and normative and (b) systematically related.

This, I believe, addresses both the conceptual and the empirical problems outline above, but only in an "in principle" fashion. What more specific evidence can be offered in support of this bridging role of bodily mimesis and mimetic schemas in human ontogeny? I wish to conclude with a rather a bold hypothesis: *autism can be (in part) attributed to a direct or indirect impairment in bodily mimesis.* What are the grounds for this suggestion?

Authorities on the still mysterious syndrome of autism, such as Firth (1989), Baron-Cohen (1995), Tager-Flusberg (2000) and Hobson (2004) differ extensively in their theoretical accounts, but they all agree that autism is a developmental disorder characterized by impairments in *social interaction, communication* and *imagination*. Bodily mimesis and

[11] I wish to thank Esa Itkonen for helping me clarify this point.

mimetic schemas have been argued to be crucially involved in all three of these capacities, so this is the first connection. More specifically, children with autism, even in the advanced range of the spectrum, do not imitate appropriately (Hobson and Lee 1999), if they do at all (Charman et al. 1997), and here the connection to bodily mimesis is obvious. Hobson (2004) presents a review of much evidence that such children have difficulties in understanding and sharing, above all *emotional* experiences with others, and attributes these problems to a fundamental disturbance in the capacity of autistic children to *identify* with others, i.e. primary intersubjectivity.

At the same time, such essentially *social* deficits in autism could also have causes that are not specifically directed at social cognition, such as *enhanced local processing* (Happé 1999) or *reduced generalization capacity* (Plaisted 2001), which have been proposed in order to explain the peculiarities of (non-social) perceptual processing in autism, including savant skills in music and mathematics, which are difficult to account assuming a primarily social deficit (Happé 1999).

In fact, the hypothesis here suggested can possibly reconcile the "social" and "non-social" explanations of autism. Not withstanding whether children with autism have a direct impairment in relevant aspects of their (innate) body schema, or fail to develop it sufficiently due to generalization difficulties, there is accumulating evidence that children who are eventually diagnosed with austim have peculiarities in their early motoric development (Teitelbaum et al. 1998; Baranek 1999; Zwaigenbaum et al. 2005). A neuroscientifically supported explanation of these difficulties has been advanced: while normally developing children use predominantly an *anticipatory* mode of control in performing voluntary actions, children with autism use a feedback mode of control which is much less efficient (Schmitz et al. 2003). The latter authors interpret this as a deficiency in the development of "action representations" in autism and pose the question whether such representations may be a prerequisite for the development of "social and cognitive representations". The present theory would, naturally, propose a positive answer, identifying the deficient "action representations" in autism with mimetic schemas, or at least with their precursors. While the disturbances may be initially (mostly) sensorimotor, they would affect negatively the mimetic mapping between exteroception and proprioception, and thus the ability to identify with others, to imitate and to develop a normal *body image.* Again, I regarded it as an unresolved empirical question whether at the bottom of this process may lay an impaired generalization capacity (Plaisted 2001) or whether different initial causes can result in a similar outcome: deficient mimetic schemas. The present focus is in the fact that with impairment in bodily mimesis, we should expect to find deficits in imagination, gesture and eventually language in autism, a prediction that is confirmed. Even high-performing children with autism who manage to acquire language more or less fluently (a relatively small proportion of all children diagnosed with the syndrome), still display not only "pragmatic" deficits, but semantic and grammatical ones (Tager-Flusberg 2000) showing that their knowledge is not completely in tune with that of their community.

In a recent publication, Tomasello et al. (2005) provide a similar account of a species-specific characteristic of human beings leading to cultural evolution, and speculate that this adaptation is impaired in autism:

> Our proposal for this "small difference that made a big difference" is an adaptation for participating in collaborative activities involving shared intentionality – which requires selection during human evolution for powerful skills of intention reading as well as motivation to share psychological states with others. In ontogeny, these two components […] intermingle from the beginning to produce a *unique developmental pathway for human cultural evolution, involving unique forms of social engagement, symbolic communication, and cognitive representation* (ibid: 690, my emphasis).

In particular, Tomasello et al (2005) claim that children with autism fail to develop (adequate) "dialogic cognitive representations". However, Tomasello et al.'s theory can be said to have two major drawbacks. First, it fails to explain how in normal children "the process of emotional engagement yields the forms of sharing that are critical for specifically human communication and thinking" as pointed out by Hobson (2005: 704). Even more critically, the nature and development of the proposed cognitive representations "which are in some way internalized in Vygotskyan fashion" (Tomasello et al. 2005: 689) remain vague. The theory presented here is, I believe, more specific in both of these points: what is internalized are mimetic schemas, and the fact that these are rooted in proto-mimesis as a form of primary intersubjectivity, shows the continuity between the "dyadic (emotional) engagements" of normal infants, the triadic engagements of older infants, and eventually the "collaborative engagements" based on shared mimetic schemas, which can be communicated first gesturally and eventually through language. In sum, bodily mimesis and mimetic schemas can help explain the nature of human cognitive specificity, being what allows non-autistic children to co-construct a world of meaning that is shared with their elders and peers.

## 6. Summary and conclusions

We began by addressing the relationship between intersubjectivity and language, and suggested relating these (in evolution and development) through the concepts of bodily mimesis and mimetic schemas. The latter were then explicated as cognitive structures which are:

- *representational* – "running" the schema (either in reality or in imagination) is differentiated from the activity, action or object to which is corresponds.
- *accessible to consciousness* – even though it is not always a matter of focal, but marginal consciousness (Gurwitsch 1964).
- *intermediately abstract* – since mimetic schemas are generalizations of bodily activities or actions, they are less abstract than "image schemas" such as PATH, but more so than concrete episodes, or specific action schemas
- *proprioceptively based* – even if the action is mentally "simulated" rather than actually re-enacted.

- *pre-reflectively shared* – since they derive from imitating culturally salient activities and actions, both their representational and experiential content can be "shared" by the members of the community, and thus constitute *cultural* representations.

Mimetic schemas have been argued to function as a pre-linguistic "ground" (and evolutionary exaptation) for language since they constitute the first true (i.e. "accessible") concepts. They also allow the realization that others have similar concepts, though possibly different perspectives. In this way, they serve as precondition for communicative intentions (in the form triadic mimesis) and at the same time provide basic semantic structure (e.g. representations of activities and actions) for the acquisition of language. Crucially, they bootstrap the acquisition of verbs, which are essential for the acquisition of grammar (Tomasello 1992, 2003). Furthermore, as a ground for public, conventional symbols, mimetic schemas can help explain how both "cognitive" (representational) and "affective" (experiential) meaning can be communicated through language, since both aspects can be – to various degrees – shared by communicators with similar bodily experiences.

At the same time, it should be held in mind that mimetic schemas do not *constitute* linguistic meanings, which are not only intersubjective, but conventional-normative and systematically interrelated in a way that mimetic schemas are not. Acquiring language and engaging in narrative practices introduces therefore yet another dimension in the development of human intersubjectivity (e.g. Hutto 2004, in press), which I have here dealt with only cursorily, since the focus has been on bridging the gap between "individual" experience and collective language.

Without claiming expertise, I have sought empirical support for the theory presented in the literature on sensory and cognitive impairments. Claiming that bodily mimesis and mimetic schemas (rather than "image schemas" or other related constructs) lie at the roots of human cognition are consistent with a number of clinical observations, such as those that deaf children are not aversely cognitively affected, and are perfectly capable of even creating a language from mimesis given the right conditions, blind children provided with rich bodily and verbal interaction will not be cognitively and linguistically retarded and even children with severe motor impairments, e.g. SMA will be relatively unaffected, suggesting that mimetic schemas are not sensorimotor structures but conscious reenactments.

On the other hand, children with autism, who are not so much mind-blind as "body blind" (Gallagher 2005) – or rather: "short-sighted" – will be most aversely affected. If the theory presented in this article can contribute to the working out of a satisfactory explanation, and perhaps even suggest novel methods of treatment, for this still mysterious impairment of human intersubjectivity, that would be its strongest theoretical and practical validation.

**Acknowledgements**

I wish to thank the editors of this volume and an anonymous reviewer as well as Pamela Heaton, Mats Andrén, Göran Sonesson, Esa Itkonen, Carrita Paradis, Svetlana Ozol,

Gabriella Rundblad, Chris Sinha and Görel Sandström for comments on an earlier version of this article.

## References

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.

Arbib, M. (2005) From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28/2: 105-124.

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.

Baars, B. J. (2003) The double life of B.F. Skinner: Inner conflict, dissociation and the scientific taboo against consciousness. *Journal of Consciousness Studies*, 10/1: 5-25.

Baldwin, D. (1995) Understanding the link between joint attention and language. In *Joint Attention: Its Origins and Role in Development*. C. Moore and P. Dunham (eds.) Hillsdale, NJ: Erlbaum.

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

Baranek, G. (1999) Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9-12 months of age. *Journal of Autism and Developmental Disorders*, 29/3: 213-224.

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

Barsalou, L. W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences*, 22: 577-660.

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

Bermúdes, J. L. (2005) *Philosophy of Psychology: A Contemporary Introduction.* New York: Routledge.

Bishop, D. (1989) Autism, Asperger's syndrome and semantic-pragmantic disorder: Where are the boundaries? *British Journal of Disorders in Communication* 24: 107-121.

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

Bowerman, M. (1996) Learning how to structure space for language–A cross-linguistic perspective. In: P. Bloom, M. A. Peterson, L. Nadel & M. Garret, eds. (1996) *Language and Space*, 385-436. Cambridge, Mass.: MIT Press.

Brinck, I. (2003) The pragmatics of imperative and declarative pointing. *Cognitive Science Quarterly,* 3(4): 429-446.

Bruner, J. (1983) *Child's Talk*, New York: Norton.

Bruner, J. (1990) *Acts of Meaning*. Cambridge, MA: Harvard University Press.

Charmam, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G. and Drew, A. (1997) Infants with autism: An investigation of empathy, pretend play, joint attention and imitation. *Developmental Psychology* 33/5: 781-789.

Choi, S. and Bowerman, M. (1991) Learning to express notion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41/1-3: 83-121.

Chomsky, N. (1975) *Reflections on Language*. New York: Pantheon Press.

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

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

Clark, H. and Marshall, C.R. (1981) Definite reference and mutual knowledge, In A. Joshi, B. Webber and I. Sag, eds., *Elements of Discourse Understanding*, Cambridge: Cambridge University Press.

Dahl, Ö. (1985) *Tense and Aspect Systems.* Oxford: Basil Blackwell.

Damasio, A. (2000) *The Feeling of What Happens*. *Body, Emotion and the Making of Consciousness.* New York: Harvester.

Davidson, D. (2001) *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.

Decety, J. and Chaminade, T. (2005) The neurophysiology of imitation and intersubjectivity. In: S. Hurley and N. Chater. (eds.) *Persepctives on Imitation: From Cognitive Neuroscience to Social Science, Vol 1. Mechanisms of Imitation and Imitation in Animals* 119-140. Cambridge, Mass: MIT Press.

Descartes, R. (1968 [1637]) *Discourse on Method*, trans. F.E. Sutcliff. Harmondsworth: Pengin Books.
Donald, M. (1991) *Origins of the Modern Mind. Three Stages in the Evolution of Culture and Cognition.* Harvard: Harvard University Press.
Donald, M. (2001) *A Mind So Rare. The Evolution of Human Consciousness*. New York: Norton.
Durst-Andersen, P. (1992) *Mental Grammar: Russian Aspect and Related Issues.* Columbus, Ohio: Slavica Publishers.
Edelman, G. (1992) *Bright air, brilliant fire. On the matter of the mind*. New York: Basic Books, HarperCollins Publications.
Elman, J. (1993) Learning and development in neural networks: the importance of starting small. *Cognition* 48: 71-99.
Fodor, J. A. (1975) *The Language of Thought*. New York: Crowell.
Fodor, J. A. (1983) *The Modularity of Mind*, Cambridge, Mass.: MIT Press.
Frith, U. (1989) A new look at language and communication in autism, *British Journal of Disorders of Communication* 24: 123 - 150.
Gallagher, S. (2005) *How the Body Shapes the Mind.* Oxford: Oxford University Press.
Gallese, V. and Lakoff, G. (2005) The brain's concepts: The role of the sensory-motor system in conceptual knowledge. Cognitive Neuropsychology 22: 445-479.
Gallese, V., Keyners, C. and Rizzolatti, G. (2004) A unifying view of the basis of social cognition. *Trends in Cognitive Sciences,* 8/9: 396 – 403.
Gardner, H. (1985) *The Mind's New Science. A History of the Cognitive Revolution*. New York: Basic Books.
Georgieff, N. and Jeannerod, M. (1998) Beyond consciousness of external events: A 'Who' system for consciousness of action and self-consciousness. *Consciousness and Cognition* 7: 465-77.
Gibson, J. J. (1979) *The Ecological Approach to Visual Perception.* Boston: Houghton-Mifflin.
Glüer, K. and Pagin, P. (2003) Meaning theory and autistic speakers. *Mind and Language*, 18/1: 23-51.
Goldman, A. (1989) Interpretation psychologized. *Mind and Language* 4: 161-185.
Grady, J. (2005) Image schemas and perception: Refining a definition. In: *From Perception to Meaning: Image Schemas in Cognitive Linguistics,* ed. Beate Hampe, 35-56. Berlin: Mouton.
Grice, H. P. (1957) Meaning. *Philosophical Review* 66: 377-88.
Gurwitsch, A. (1964) *The Field of Consciousness*. Pittsburgh: Duquesne University Press.
Hampe, B., ed. (2005) *From Perception to Meaning: Image Schemas in Cognitive Linguistics.* Berlin: Mouton.
Happé, F. (1999) Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences* 3/6: 216-222.
Hobson, P. (2004) *The Cradle of Thought: Exploring the Origins of Thinking*. London: Pan.
Hobson, P. (2005) The interpersonal foundations of thinking. Commentary to Tomasello et al. *Behavioral and Brain Sciences* 28: 703-704.
Hobson, P. and Lee, A. (2001) Imitation and identification in autism. *Journal of Child Psychology and Psychiatry* 40/4: 649-659.
Hobson, P., Lee, A. and Brown, R. (1999) Autism and congenital blindness. *Journal of Autism and Developmental Disorders*, 29/1: 45-56.
Husserl, E. (1999) [1907] *The Idea of Phenomenology,* trans. L. Hardy. Collected Works, Vol 8. Dordrecht: Klewer.
Hutto, D. (2004) The limits of spectatorial folk psychology. *Mind and Language* 19: 548-73.
Hutto, D. (in press) *Folk Psychological Narratives: The Socio-cultural Basis of Understanding Reasons*, Cambridge MA: Cambridge University Press.
Iacoboni, M. (2005). Understanding other: Imitation, language and empathy. In S. Hurley, & N. Chater, *Perspectives on Imitation: From Neuroscience to Social Science. Vol 1. Mechanisms of Imitation and Imitation in Animals* (pp. 77-100). Cambridge, Mass.: The MIT Press.
Itkonen, E. (1978) *Grammatical Theory and Metascience: A Critical Inquiry into the Philosophical and Methodological Foundations of "Autonomous" Linguistics* Amsterdam: Benjamins.
Itkonen, E. (1997) The social ontology of meaning. *SKY Journal of Linguistics 10, 49-80.*
Itkonen, E. (2003) *What is Language? A Study in the Philosophy of Linguistics. University of Turku:* Publications in General Linguistics 6.
Jackendoff, R. (1983) *Semantics and Cognition*. Cambridge, Mass.: MIT Press.

Jeannerod, M. (1997) *The Cognitive Neuroscience of Action.* Cambridge, MA: Blackwell Science.
Johnson, M. (1987) *The Body in the Mind*. Chicago: University of Chicago Press.
Johnson, M. (2005) The philosophical significance of image schemas. In: *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, ed. Beate Hampe, 15-33. Berlin: Mouton.
Johnson, M. and Lakoff, G. (2002) Why cognitive linguistics requires embodied realism. *Cognitive Linguistics* 13 (3): 245-263.
Johnson, M. and Rohrer, T. (2007) We are live creatures: Embodiment, pragmatism and the cognitive organism. In: T. Ziemke, J. Zlatev, and R. Frank (eds.) *Body, Language and Mind. Vol 1. Embodiment*. Berlin: Mouton.
Kita, S. (ed.) (2003) *Pointing: Where language, culture and cognition meet.* Mahwah, N.J.: Erlbaum.
Lakoff, G. (1987) *Women, Fire and Dangerous Things: What Categories Reveal about the Mind.* Chicago: University of Chicago Press.
Lakoff, G. and M. Johnson (1999) *Philosophy in the Flesh*: *The Embodied Mind and its Challenge to Western Thought.* New York: Basic Books.
Landau, B. and Gleitman, L. (1985) *Language and Experience. Evidence from the Blind Child*. Cambridge, Mass.: Harvard University Press.
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.
Leibniz, G. W. (1973) [1787] *Monadology*, in M. Morris and G.H.R. Parkinson (trans.) Philosophical Writings. London.
Lewis, D. K. (1969) *Convention: A Philosophical Study.* Cambridge MA: Harvard University Press.
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.
Lutz, A. and Thompson, E. (2003) Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies* 10/9-10: 31-52.
Mandler, J. (2004) *The Foundations of Mind: Origins of Conceptual Thought*. Oxford: Oxford University Press.
Maslin, K. T. (2001) *An Introduction to the Philosophy of Mind.* Cambridge: Polity Press.
Meltzoff, A. and Gopnik, A. (1993) The role of imitation in understanding other persons and developing a theory of mind. In: S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen (eds.) *Understanding other Minds: Perspectives from Autism*. Oxford: Oxford University Press.
Meltzoff, A. and Moore, M. K. (1977) Imitation of facial and manual gestures by human neonates. *Science* 198: 75-78.
Meltzoff, A. and Moore, M. K. (1983) Newborn infants imitate adult facial gestures. *Child Development,* 54: 702-709.
Menyuk, P. and Quill, K. (1985) Semantic problems in autistic children. In E. Schoper and G. Mesibov (eds.) *Communication Problems in Autism.* New York: Plenum Press.
Merleau-Ponty, M. (1979) [1945] *Phenomenology of Perception*. London: Routledge & Kegan Paul.
Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M. and Matsuzawa, T. (2004) Imitation in neonatal chimpanzees (*Pan troglodytes*). *Developmental Science*, 7/4: 437-42.
Nelson, K. (1996) *Language in Cognitive Development. The Emergence of the Mediated Mind*. Cambridge: Cambridge University Press.
Nelson, K. and Shaw, L. K. (2002) Developing a socially shared symbolic system. In: J. Byrnes and E. Amseli (eds.) *Language, Literacy and Cognitive Development,* 27-57. Mahwah, N.J.: Erlbaum.
Newton, N. (1996) *Foundations of Understanding.* Amsterdam: Benjamins.
Newton, N. (2003) Representation in theories of embodied cognition. *Theoria et Historia Scientarum* 7/1: 181-193.
Piaget, J. (1951 [1945]) *Play, Dreams and Imitation in Children.* Translated from *La Formation du Symbole Chez l'enfant*. Delachaux et Niestlé.
Piaget, J. (1954) *The Construction of Reality in the Child*. New York: Basic Books
Pinker, S. (1994) *The Language Instinct*. New York: William Morrow.

Plaisted, K. C. (2001) Reduced generalization in autism: An alternative to weak central coherence. In J.A. Burack, T. Chapman, N. Yimiya, and P.R. Zelazo (eds.) *The Development of Autism*, 149-169. Mahwah NJ: Lawrence Erlbaum.

Plunkett, K. and Virginia M. (1993) From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48/1: 21-69.

Plunkett, K., Sinha, C., Møller, M. and Strandsby, O. (1992) Symbol grounding or the emergence of symbols? Vocabulary growth in children and in a connectionist net. *Connection Science* 4/3-4: 293-312.

Putnam, H. (1975) The meaning of "meaning". In: *Mind, Language and Reality.* Cambridge: Cambridge University Press.

Regier, T. (1996) *The Human Semantic Potential: Spatial Language and Constrained Connectionism*. Cambridge, Mass.: MIT Press.

Riviére, J. and Lècuyer, R. (2003) The C-not-B error: a comparative study. *Cognitive Development* 18: 285-297.

Riviére, J. and Lècuyer, R. (2002) Spatial cognition in young children with spinal muscular atrophy. *Developmental Neuropsychology* 13: 165-184.

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

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

Ryle, G. (1949) *The Concept of Mind.* London: Hutchingson.

Sapir, E. (1928) The unconscious patterning of behaviour in society. In E.S. Dummer (ed.) *The Unconscious*, 114-142. New York: Knopf.

Scheler, M. (1954 [1912]) *The Nature of Sympathy*, trans. P. Heath. London: Routledge and K. Paul.

Schmitz, C, Martineau, J, Barthelemy, C, Assaiante, C. (2003) Motor control and children with autism: deficit of anticipatory function. *Neuroscience Letter* 348: 17-20.

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

Searle, J. (1999) *Mind, Language and Society. Philosophy in the Real World*. London: Weidenfeld and Nicolson.

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*. New York: Harverster Press.

Sinha, C. (1999) Grounding, mapping and acts of meaning. In: Theo Janssen and Gisela Redeker (eds.) Cognitive Linguistics: Foundations, Scope and Methodology, 223-255. Berlin & New York: Mouton.

Skinner, B. F. (1953) *Science and Human Behavior.* New York: Macmillian.

Sonesson, G. (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.

Sonesson, G. (2007) From the meaning of embodiment to the embodiment of meaning: A study in phenomenological semiotics. In: T. Ziemke, J. Zlatev, and R. Frank (eds.) *Body, Language and Mind. Vol 1. Embodiment,* 85-128. Berlin: Mouton de Gruyter.

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

Stern, D. N. (1985) *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology.* New York: Basic Books

Tager-Flusberg, H. (2000) Language and understanding minds: Connections in autism. In: S. Baron-Cohen, H. Tager-Flusberg, D. Cohen (eds.) *Understanding Other Minds: Perspectives from Developmental Neuroscience*, 124-149. Oxford: Oxford University Press.

Talmy, L. (2000) *Toward a Cognitive Semantics*, *Vol I and Vol II.* Cambridge, Mass.: MIT Press.

Taub, S. (2005) The Symbolic Base: Analyzing the Blend Structure of Iconic Signs. Paper presented at the conference *New Directions in Cognitive Linguistics*, Brighton 23-25 October 2005.

Teitelbaum, P., Teitelbaum, O., Nye J., Fryman, J. and Mauer R.G. (1998) Movement analysis in infancy may be useful for early diagnosis of autism. Proceedings of the National Academy of Sciences 95: 13982-7.

Thompson, E. (2001) Empathy and consciousness, *Journal of Consciousness Studies* 8 (5-7): 1-32.

Tomasello, M. (1992) *First verbs: A Case Study of Early Grammatical Development. New York:* Cambridge University Press.
Tomasello, M. (1999) *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
Tomasello, M. (2003) *Constructing a Language: A Usage-based Theory of Language Acquisition.* Cambridge, MA: Harvard University Press.
Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005) Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences 28: 675-691.*
Trevarthen, C. (1979) Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (ed.) *Before Speech*. Cambridge: Cambridge University Press
Varela, F., Thompson, E. and Rosch, E. (1991) *The Embodied Mind. Cognitive Science and Human Experience.* Cambridge, Mass.: MIT Press.
Vygotsky, L. S. (1962) [1934] *Thought and language*. Cambridge, Mass.: MIT Press.
Vygotsky, L. S. (1978) *Mind in Society. The Development of Higher Psychological Processes.* Cambridge, Mass.: Harvard University Press.
Watt, R. (1991) *Understanding Vision.* New York: Academic Press.
Whiten, A., Horner, V., Litchfield, C. A. and Marshall-Pescini, S. (2004) How do apes ape? *Learning and Behaviour*, 32/1: 36-52.
Wittgenstein, L. (1953) *Philosophical Investigations*. Oxford: Basil Blackwell.
Zelazo, P. D. and Lourenco, S. F. (2003) Imitation and the dialectic of representation. *Developmental Review* 23: 55-78.
Zlatev, J. (1999) Situated embodied semantics and connectionist modeling, In J. Allwood and P. Gärdenfors (eds.) *Cognitive Semantics: Meaning and Cognition*, 173-194. Amsterdam: Benjamins.
Zlatev, J. (2000) The mimetic origins of self-conciousness: in phylo-, onto- and robotogenesis, *Proceedings of the Third Asia-Pacific Conference on Simulated Evolution And Learning (SEAL2000),* Nagoya, Japan.
Zlatev, J. (2002) Mimesis: The "missing link" between signals and symbols in phylogeny and ontogeny. In A. Pajunen (ed.) *Mimesis, Sign and the Evolution of Language,* 93-122. Publications in General Linguistics 3. University of Turku Press.
Zlatev, J. (2003) Meaning = Life (+ Culture). An outline of a unified biocultural theory of meaning. *Evolution of Communication,* 4/2: 253-296.
Zlatev, J. (2005) 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, 313-342. Berlin: Mouton.
Zlatev, J. (2007) Language, embodiment and mimesis. In: T. Ziemke, J. Zlatev, and R. Frank (eds.) *Body, Language and Mind. Vol 1. Embodiment*, 297-338. Berlin: Mouton de Gruyter.
Zlatev, J. (in press) The co-evolution of intersubjectivity and bodily mimesis. In J. Zlatev, T. Racine, C. Sinha and E. Itkonen (eds.) *The Shared Mind: Perspectives on Intersubjectivity.* Amsterdam: Benjamins.
Zlatev, J., Persson, T. and Gärdenfors, P. (2005) Bodily mimesis as "the missing link" in human cognitive evolution. *Lund University Cognitive Studies 121*.
Zlatev, J., Racine, T., Sinha, C., and Itkonen, E. (in press) *The Shared Mind: Perspectives on Intersubjectivity.* Amsterdam: Benjamins.