

The Mimesis Hierarchy of semiotic development: Five stages of intersubjectivity in children

Jordan Zlatev

Department of Linguistics, Centre for Languages and Literature

Centre for Cognitive Semiotics

Lund University, Sweden

Abstract

The paper proposes that intersubjectivity develops in children along a progression of five, more or less distinct, stages of semiotic development. The theoretical model within which this is couched is the Mimesis Hierarchy (MH) model (Zlatev & Andr n 2009). As in previous treatments, the MH-model focuses on bodily mimesis, its “precursors” (empathetic perception) and “post-developments” (conventionality, language and narrative). Mimesis is pivotal since it provides the basis for the development of (i) conventions (through imitation), (ii) intentional communication, and (iii) for bringing the two together in communicative, shared representations (signs). The main difference from previous applications is in the treatment of the concepts of *representation* and *communicative intent*. Due to recent empirical findings, and a more bodily-enactive and social-oriented perspective, I propose that Stage 2 gives rise to imitation and *mimetic schemas* (Zlatev 2007, in press), but that the first gestures (or vocalizations) of children are neither externalizations of these “internal representations”, nor fully-fledged representations/signs on their own, but action schemas bi-directionally *associated* with particular contexts. That would explain why the onset of intentional communication occurs in Stage 3 with pointing and other deictic gestures (such as showing), which are not representations or fully-fledged (explicit) signs, but rather performative communicative acts, accompanied with makers of communicative intent. It is first in Stage 4 that the proto-representations of Stage 2 and the communicative intent of Stage 3 are combined to give rise to communicative iconic gestures, and more generally to the “insight” of using communicative, shared representations, or what is commonly referred to as symbols or signs.

1. Introduction

Children develop from birth, and possibly even earlier, not just *cognitively*, i.e. what they *know* about their surrounding physical and social environment, but in terms of *meaning*, i.e. their value-based relationship to the world as subjects of experience (Zlatev 2009). With time, this relationship changes, acquires new dimensions and undergoes transitions. In other words: children undergo *semiotic development*.¹

Different theorists have focused on different aspects and periods of such development. Trevarthen (1979) and Trevarthen & Hubley (1978) charted basic social capacities: from neonatal imitation, to “proto-conversations” and eventually triadic interactions around objects, and described the changes during the first year of life as a gradual shift from primary to secondary intersubjectivity. From a similar theoretical approach, Reddy (2003, 2005) reported evidence for surprisingly early awareness of self and other. Moro (2011) has rather focused on the role of interactions with cultural artifacts, and how children thereby expand their semiotic horizons through the help of others. Tomasello’s (1999, 2003) interest has rather been on the second year of life, with the development of joint attention, pointing, the understanding of communicative intent, and the first indisputable steps in the acquisition of language: from the production of the first words around 14 months, through the “vocabulary spurt” around 18-20 months, to the first multi-word constructions. Nelson (1996, 2003) has convincingly shown how the development of language implies cognitive-semiotic development, in particular through the ability to construct narratives, and with their help autobiographical memories, from the fourth year of life. But while language is a key semiotic resource, from a cognitive-semiotic perspective, other resources should not be underestimated. Thus, the studies of DeLoache (2004) on children’s progressive understanding of pictorial representation constitute an important complement.

Such research has given us important insights on children’s semiotic development. One drawback, however, which becomes evident if we compare any of the mentioned studies with Piaget’s classical developmental theory (Piaget 1954, 1962), is that they appear quite specific, with focus on particular ages and cognitive-semiotic skills such as interpersonal interactions, artifacts, intentions, words, narratives, pictures... While some would claim that the quest for such general developmental theories as Piaget’s is outmoded, there is something quite unsatisfactory with the current “particularist” zeitgeist in much of developmental psychology. Lenninger (2012), for example, urges to consider children’s semiotic development more holistically.

The particularist flavor of most studies of semiotic development stems, in part, from the fact that the mastery of specific semiotic resources such as language and picture-understanding takes place at

¹ Of course, cognition and meaning are closely intertwined, and the term “semiotic development” should be viewed as a short-hand for cognitive-semiotic development.

different periods of development, and it has not been clear whether, and if so *how* these are related. In this respect a crucial social-semiotic skill, which was intentionally omitted above, differs: *children's gestures* which are generally agreed to co-develop with speech (Bates et al. 1979; Iverson & Goldin-Meadow 1998; McNeill 2005; Andrén 2010). Gesture, however, has been argued by Donald (1991, 2001) to be part of a more general cognitive-semiotic suite, for which Donald reserves the Aristotelian concept *mimesis*, understood as “the ability to produce conscious, self-initiated, representational acts that are intentional but not [narrowly] linguistic” (Donald 1991: 168). According to Donald, bodily *mimesis* evolved in our ancestors during the past two million years, as “the result of evolving better conscious control over action. In its purest form, it is epitomized by four uniquely human abilities: mime, imitation, skill [rehearsal], and gesture.” (Donald 2001: 263)

In previous work, I have argued that bodily *mimesis* is intimately linked with the human capacity for *intersubjectivity*, understood as “the sharing of affective, perceptual and reflective experiences between two or more subjects, [which] can take different forms, some more immediate, while others more mediated by higher cognitive [-semiotic] processes” (Zlatev 2008a: 215). Furthermore, since it is possible to identify close “precursors” to *mimesis* on the one hand, and language can be seen as essentially post-mimetic on the other, I have proposed a *Mimesis Hierarchy*, consisting of five more or less distinct levels, each building cumulatively on top of the previous. The application of this model to human cognitive-semiotic evolution, and in particular to the evolution of language, has been productive (Zlatev 2008b). Since the levels of the *Mimesis Hierarchy* are sufficiently generally defined (see Section 2), it is also possible to apply the model to children’s semiotic development, without evoking any simplistic notion of “recapitulation”. This was essentially the argument presented in previous work (Zlatev & Andrén 2009), where we focused on the development of children’s gestures, and to some extent speech.

The goal of the present article is to elaborate on this, proposing that the five levels of the model correspond to *five more or less distinct stages in the development of intersubjectivity*: from basic empathy to folk psychology. Since intersubjectivity is arguably one of the essential characteristics of the human mind (Zlatev et al. 2008), these stages should also be expected to involve other cognitive-semiotic skills. Thus, I submit that the *Mimesis Hierarchy* may serve as the basis for *a general model of semiotic development*, unifying many of the approaches mentioned earlier. A likely objection to a multi-stage developmental model was anticipated by Zlatev & Andrén (2009: 380-381):

The concept of *developmental stage* played a central role in nearly all the classic theories of cognitive, emotional, and moral development of the past century, such as those of Montessori, Piaget, Kohlberg, Freud, Erikson and Vygotsky. In language acquisition, “it is possibly the most often used term” (Ingram 1989: 32). During the last two decades, however, the stage concept has come under a good deal of critique for being inconsistently defined (or not defined at all), failing to predict the varying performance of children in different cognitive domains (Gardner 1992),

being too discrete and static (Siegler 1996) and often implying a complete replacement and “dismantling” of the previous stage, while “no emerging domain disappears; each remains active and interacts dynamically with all the others” (Stern 1998: xii). However, such critiques can be taken as implying the need to *improve* on the notion of developmental stage, rather than reject it.

The concept assumed in the present article is similar to that proposed earlier: *a stage in the development of X, is a (relatively stable) period in life, characterized by the consolidation of a novel cognitive-semiotic capacity, which may dominate the expression of X at this stage, but does not replace capacities from previous stages.* Since “modularity” is no longer as generally accepted as it was in the last decades of the past century, such a concept may be (once again) found fruitful. Section 2 presents the concept of bodily mimesis, and the Mimesis Hierarchy in its application to *the development of intersubjectivity in children.* In Section 3, I will briefly review research that mostly supports the original model (Zlatev & Andr n 2009), but also calls for some important modifications. These will be summarized in the final section, which also provides brief comparisons with similar models, and general conclusions.

2. Bodily mimesis and the Mimesis Hierarchy model of semiotic development

Bodily mimesis is either realized through action, or else this action could be imagined, virtual, or as currently popularly phrased: “simulated”. To delineate it from other similar phenomena such as mimicry (“from below”), or signed language (“from above”), the following definition will suffice:

(DEF) An act of cognition or communication is **an act of bodily mimesis** if and only if:

- 1) It involves a *cross-modal mapping* between exteroception (e.g. vision) and proprioception (e.g. kinesthesia).
- 2) It is *under conscious control* and is *perceived by the subject to be similar* to some other action, object or event.
- 3) The subject *intends* the act to stand for some action, object or event *for an addressee*, and for the addressee to recognize this intention.
- 4) It is *not fully conventional* (and normative).
- 5) It does *not divide (semi)compositionally* into meaningful sub-acts that *systematically relate* to other similar acts (as in grammar).

This is nearly the same definition as that provided earlier (Zlatev 2008a, 2008b), with the difference that clause (2) has been simplified, and now explicitly involves *similarity*: as in acts of imitation, or in bodily-iconic signs (gestures, pantomimes). Unlike in previous treatments, acts of pointing qualify as mimetic acts to the extent that they are imitated, but not in general. However, since the specifics of pointing acts are largely culture-typical, and there is good evidence that they are learned by children at least in part through imitation (Tomasello 1999), pointing should nevertheless be regarded as an instance of bodily mimesis, and when accompanied with communicative intent, as triadic mimesis (see below).

The Mimesis Hierarchy (hence, MH) follows straightforwardly from this definition, once it is stated that if only (1) is fulfilled, the act is one of *Proto-mimesis*; that (1) and (2) together qualify for *Dyadic (non-intentionally communicative) mimesis*, and only when (3) is added is there full *Triadic (intentionally communicative) mimesis*. When also (4) and (5), the negative criteria in the definition, are fulfilled we have rather two post-mimetic stages: *Protolanguage* (with little systematicity, i.e. “grammar”) and *Language*, with sufficient systematicity to allow the construction of discourse and narratives. Table 1 displays the MH applied to the development of intersubjectivity, listing (i) crucial novel cognitive-semiotic capacities that define the stage compared to its predecessor, (ii) example skills that may be regarded as “behavioral indexes” and (iii) approximate age-periods. All of these will be further motivated and illustrated in the following section.

Prior to that, I wish to highlight three points. First, the “novel features” for each successive stage are formulated in a way that expresses their fundamentally interpersonal character, with some changes compared to earlier formulations, especially concerning Stages 3 and 4 and the transition between them. Second, many of the aspects of semiotic development mentioned in Section 1, from neonatal imitation to narrative, figure as specific “skills” in the model: a testimony to its integrating character. Third, the MH is a “layered model” in the sense of Stern (1998) or the “Russian doll” model of empathy of de Waal (2007), where higher levels engulf lower ones rather than replace them, as in the classical Piagetian framework (at least as commonly interpreted).

Table 1. The Mimesis Hierarchy of children’s semiotic development, with focus on the development of intersubjectivity

Stage	Novel capacity	Examples of cognitive-semiotic skills	Approximate age
1	Proto-mimesis	Empathetic perception	0-9 m
2	Dyadic mimesis	Volitional control and Imitation	9-14 m
3	Triadic mimesis	Communicative intent	14-20 m
4	Protolanguage	Communicative, conventional representations (“signs”)	20-30 m
5	Language	Language-mediated folk psychology	30 m -

3. Five stages in the development of intersubjectivity

The challenges for any stage model of development is to provide (i) an account of the factors that organize the *coherence* of a particular stage, (ii) link these which particular *manifestations*, as testified by evidence and (iii) account for the factors (other than maturation) bringing about a *transition* of a consecutive stage. The aim of this section is to provide (i)-(iii), albeit in summary fashion, for the each of the five stages.

Stage 1: Empathetic perception (0-9 months)

The phenomenological tradition, and prominently Merleau-Ponty (1962), has contributed to a notion of perception as active and empathetic, in which the feeling body (*Leib*) “resonates” with the world, and especially with con-specifics. Despite some exaggerated initial enthusiasm, the “mirror neuron” literature of the past decade (cf. Iacoboni 2008 for a summary) has provided a series of hard-science confirmations of this conception, according to which in perception, the actions of others are “mapped” onto one’s own bodily actions and sensations.

The now classical studies of *neonatal imitation* of Meltzoff and Moore (1977, 1983), showing that newborn babies are capable of imitating simple movements involving mouth-opening, tongue-protrusion, lip-protrusion, and simple hand movements, have provided evidence that at least some of this capacity is innate, i.e. present at birth (Gallagher 2005). At the same time, it undoubtedly

undergoes gradual post-natal development, e.g. as caregivers engage in “imitating games”, e.g. matching the baby’s first spontaneous smiles with their own. Infants thus spontaneously learn to share in the somatosensory states of others, and thus realize a basic form of empathy, which may generally be defined as “any process where the attended perception of the object’s [i.e. other’s] state generates a state in the subject that is more applicable to the object’s [i.e. other’s] state or situation than to the subject’s own prior state or situation” (Preston & de Waal 2002: 4). In its simplest form, this can be observed in *emotional contagion*, familiar from situations in which crying “catches on” in a post-natal ward. Towards the 6th month, infants also learn to orient themselves in the direction where the other is looking: a form of *attention contagion*, or the simplest kind of joint attention, “synchronous”, (Zlatev, Brinck & Andrén 2008, see below)

While the proto-mimetic behaviors mentioned in the previous paragraph have also been demonstrated (albeit in weaker forms) in other non-human primates (Preston & de Waal 2002; Zlatev 2008a), other more finely tuned “orchestrations” of what Trevarthen (1979) refers to as *primary intersubjectivity* have not. For example, starting from 2 months, the “proto-conversations” of caregivers and infants take on the quality of a rhythmic “dance”, and frustration follows if this attunement is disrupted. Working in the same tradition, Reddy has argued that starting from several months infants “show an awareness of others as attending beings, as well as an awareness of self as an object of others’ attention” (Reddy 2003: 357), displayed in phenomena such as mutual gaze, intense smiling, coyness, “calling” vocalizations, showing-off etc. The range of such primary intersubjective engagements shows that “mirroring” or “exteroception-proprioception matching” is not sufficient to characterize what is here called empathetic perception: it should also involve (i) spontaneous anticipations (called *protentions* in the phenomenological literature), (ii) responses across different modalities (cross-modality), and (iii) at least some degree of self-other differentiation. Still, even such enriched empathetic perception does not require full voluntary control of one’s movements, nor an explicit distinction between self and other; as noted by Reddy (2003: 401): “older infants reveal a greater focus on the self and the younger ones reveal a *more immersed, less detached* focus on the other”.

Mutual gaze has been regarded by Reddy and others as a powerful index of primary intersubjectivity, and has been suggested to be human-specific. On the other hand, Bard et al. (2005) have provided evidence for similar rates and durations of mutual gazing in parent-infants dyads among human beings and chimpanzees. To throw light on this, we conducted a comparative study in which 5 ape dyads (three chimpanzee, one bonobo, one gorilla) and 5 human dyads (living in Lund, Sweden) were recorded for 3 hours per dyad during typical interactions (mean age for apes = 8;26, children = 6;11). Indeed, we found that the rates of mutual gaze for the human dyads were much higher (34.9 vs. 1.8 per hour) and of much longer duration (3.33 vs. 0.94 sec.). Obviously, these differences cannot be

attributed only to biological factors since the infants were being raised in radically different environments and cultures. However, it underscores the importance of the “gaze of the other”, and the meeting of minds in acts of perception, prior to the development of full motoric control, and in consort with that, a full sense of agency and “ownership” of the body (Gallagher 2005).

Stage 2: Volitional control and imitation (9-14 months)

There is considerable agreement that a transition in cognitive-semiotic development occurs around 9 months, though views on how to explain it vary considerably. As mentioned in Section 1, for Trevarthen & Hubley (1978) this marks the onset of secondary intersubjectivity, involving triangulations between infant, adult and an external object. However, joint activities with objects are observed, at least in some cultures, much earlier (Rodríguez & Moro 2008). Tomasello (1995, 1999) famously proposed that a major “social-cognitive revolution” occurs at this particular age:

At about 9 months of age, infants begin to behave in a number of ways that demonstrate their growing awareness of how other persons work as psychological beings. They look where adults are looking (joint attention), they look to see how adults are feeling toward a novel person or object (social referencing), and they do what adults are doing with a novel object (imitation learning). ... Infants also at this time first direct intentional communicative gestures to adults, indicating an expectation that adults are causal agents who can make things happen. (Tomasello 1995: 175)

However, this synchrony of developmental landmarks has been questioned. Reddy (2005) points out that infants display the marks of “understanding attention”, in particular with respect to themselves, much earlier (see Stage 1) and that “social referencing” is generally accepted to begin from 7 months. As for “intentional communicative gestures”, in the quotation, Tomasello seems to blur the distinction between (a) gestures performed intentionally (i.e. volitionally), and serving a communicative purpose though not intended as such (such as an arm stretched out in the direction of a desired object) and (b) gestures accompanied with marks of communicative intent, especially those performed for the sake of informing an addressee. While the first do indeed commence around this period, both production and comprehension of the latter will require a further stage in the development of intersubjectivity.

Still, what remains intact from the Tomasello quotation (“growing awareness”, “imitation learning”, volition) is consistent with the explanation offered by the MH-model for the transition to Stage 2 of child intersubjectivity (cf. Zlatev & Andrén 2009): namely, that what gradually develops, in interaction with others, during the first stage is a “sense of a core self” (Stern 1998) in which the body is felt to be “one’s own” and under *volitional control*. This allows a much more precise and flexible

form of imitation than that of neonates, and this imitation on its part allows a fuller understanding of the other “as a psychological being” – and vice-versa. Or as stated famously by Baldwin (1894): “My sense of myself grows by imitation of you and my sense of yourself grows in terms of myself”. But unlike the direct empathy of Stage 1, this loop of volition and imitation does not decrease, but on the contrary, *increases the awareness of a distinction between self and other*: it highlights the lack of direct control of others’ actions, and along with that motivates attempts to influence them to perform actions that are desired. This can possibly explain the surge in this stage of communicative signals, including gestures such as “imperative pointing”, which however, are not yet fully developed acts of intentional communication.

Concerning evidence for this interpretation of Stage 2, the studies performed by Mandler (2004) with infants during this age period show that they are not only capable of direct, “sensorimotor imitation” (Piaget 1962), but also of *generalized imitation*, in which infants first observe pretend-actions such as giving a sip of water to toy-objects such as airplanes, birds, jeeps and dogs, and then are given the chance to imitate with either the same object or novel ones. Mandler showed that starting from 9 months, and progressing up to 14 months, infants do not imitate “inappropriate” actions (e.g. water is given to animals, but not vehicles) and when given a novel object, do not generalize on the basis of shape (e.g. from bird to airplane), but stay within the global category (animate vs. inanimate). Further studies show that infants, again from 9 months “begin to be able to reproduce event sequences after a delay” (ibid: 232), i.e. *deferred imitation*, or the second step in the development of imitation according to Piaget, though quite a bit earlier than he anticipated. If infants are also capable of the third step: *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 1962: 279) – during this stage is not clear, and if so, the (largely) preverbal children of 9-14 months would be fulfilling two of the mimetic skills singled out by Donald (2001, see Section 1): imitation and skill-rehearsal.

What about the remaining two: “gesture” and “mime”? As noted earlier, “imperative gestures” clearly appear from 9 months, but even with gaze alternation between desired object and other person, they are relatively poor indicators of communicative intentions, since they can be learned as behavioural sequences (Brinck 2003). Declarative gestures, on the other hand, clearly indicate that the infant interacts with the other as a subject, rather than as a means-to-an-end (Tomasello’s “causal agent”). Summarizing a number of (sometimes contradictory) research findings, Carpenter, Nagell & Tomasello (1998: 20) state that “whereas declarative showing and pointing (with gaze alternation) first appear at around 9-10 months of age, they do not occur with great frequency until 12-15 months of age”, whereas imperatives do not display such a pattern of later increase.

In a study comparing types of joint attention (or “perceptual intersubjectivity”) of 12-month and 18-month-old Swedish and Thai children in naturally occurring interactions, we distinguished between three kinds of joint attention: *synchronous* (with no referential behavior or gaze alternation on the part of the child), *coordinated* (when one or both of these behaviors were present), and *reciprocal* (when in addition there was a bout of mutual gaze, confirming that the target had been mutually attended) (cf. Zlatev, Brinck & Andr n 2008). The results showed that while all three types were present in both age groups, there was a clear shift in the direction of the two latter types (coordinated and reciprocal) in the older age-group. Together with findings that the synchronous type of joint attention was also the only type present in chimpanzee dyads, this supports its interpretation as an essentially proto-mimetic phenomenon (see Stage 1), and that while the understanding of the other as a subject of experience whose attention may be influenced by deictic gestures begins at Stage 2, this understanding is not yet stabilized.

The other major type of gesture that children begin to use during this stage are conventional, from the standpoint of the community, gestures such as BYE and HEAD-NOD. However, these are highly restricted in number. Hence, it was quite surprising when Acredolo & Goodwyn (1988) showed that starting from 9 months, infants are capable of learning many so-called “baby signs”. All of these involve some action associated with a particular object or sensation: “With encouragement from parents, babies can learn to associate dozens and dozens of gestures with specific things-like flapping arms for *bird*, smacking lips for *fish*, blowing for *hot*, or even patting the chest for *afraid*.” (Acredolo and Goodwyn 2000: 84). What this finding indicates, once more, is *mimesis as imitation*. However, since there is no indication that children at this age are aware of either the conventional (mutually known) status of gestures, or their representational, iconic character (Namy, Campbell & Tomasello 2004, see below), these gestures cannot be yet regarded as either *iconic* or *symbolic signs*. In the best case, they may be seen as spatio-temporally associated *indexes*, though their referential (directed), as opposed to purely associative character would need to be established. Therefore it may be concluded that mimesis in the sense of “mime” (as in symbolic play), as well as the communicative intent necessary for “triadic mimesis”, are still absent in children around the age of one.

Stage 3: Communicative intent (14 – 20 months)

What heralds the onset of Stage 3 is precisely the understanding of *communicative intent*, as a participant in acts of intentional communication, in both production and comprehension. The notion stems from Grice (1957), according to whom to mean something by uttering/performing X is approximately equivalent to intending X to (a) produce some effect on another individual and (b) for this individual to recognize that one is intending (a). Theorists who have employed the notion (Sperber & Wilson 1995; Zlatev 2008a; Moore under review) differ in the interpretation of how complex, or

how “mentalist” it should be, but there is general agreement that communicative intent implies *at least a second-order intention (b) to recognize the primary intention (a)*.

It has been recently suggested independently by Andrén (2010) and Moore (under review), that communicative intent and *semiotic vehicles* (such as gestures, words, or pictures) can be considered independent dimensions, though intermixing in a single communicative act. Any act performed with deliberate expressiveness for the sake of an addressee will be likely understood as intentionally communicative, irrespectively of whether it “stands for” something or not (Sperber & Wilson 1995). In this respect, ostensive *mutual gaze* with an addressee can “enact” communicative intent, and specifically the second-order intention in a Gricean analysis (cf. Moore under review). On the other hand, a particular performance can function as a sign without there being a communicative intention, as when a child engages in symbolic play without anyone else present.²

We can adduce a number of recent studies in support for the proposal that (higher-order) communicative intent (and thus: triadic mimesis, see clause (3) in the definition, Section 2) begins to characterize the cognitive-semiotic performances of children in the middle of the second year of life. First, it could be remembered that the more advanced forms of joint attention (coordinated and reciprocal), which predominated in the bouts of attention sharing of 18 month-old-children as opposed to 12-month-old children, where characterized precisely by “enacting communicative intent” in the form of gaze-alternation and mutual gaze (Zlatev, Brinck & Andrén 2008).

Experimentally, the most common paradigm for assessing communicative intent is “the object-choice task”, in which an experimenter hides a reward under one of two (or three) different boxes and then he (or another experimenter, a “helper”) communicates the location of the reward by various semiotic vehicles and means of indicating communicative intent. Behne et al. (2005) showed that 14-month old children could solve the object-choice task when the experimenter pointed to the correct box, gaze-shifting between the box and the addressee, but not when pointing to the box while looking elsewhere. Ostensive gazing alone often led to finding the reward, though 24-month-old children performed better than 14- and 18-month-olds. Tomasello et al. (1997) showed similar results for the three vehicle types Point, Marker and Replica for 30 and 36-month old children, but it has not been reported how children who are two years and younger perform with other semiotic vehicles than Ostensive gaze and Point. In a recent study (Zlatev et al., under review), we included a fourth vehicle (Picture), and conducted the object-choice task with three groups of children: of 18, 24 and 30 months of age. The results were that while the 18-month olds were clearly above chance with Pointing and Marker, and the 24-month-olds

² Semiotic theories tend to privilege the role of vehicles (often regarded as “signs”, in a general sense of the term), while Gricean (and psychological) approaches tend to focus on intentions. The cognitive-semiotic approach here adopted suggests that both are non-reducible to one another though closely interacting, aspects of meaning.

were even better, only the 30-month-olds performed reliably with Picture and Replica, though about 50% of the children still failed the criterion of 5 out of 6 correct choices.

Table 2 shows the semiotic properties of the various vehicles used in our and others' studies. Since, as can be seen, the vehicles differed in terms of a number of properties, we cannot provide a definite explanation of this difference. Still, given all available research, the most likely interpretation is that while children at 18 months do not yet understand (iconic) representations like pictures and replicas (of the correct box) as communicative signs, they understand communicative intent, and do so not only for familiar vehicles such as pointing, but also novel ones as markers.

Table 2. Classifying semiotic vehicles used in object-choice studies according to the factors bodily means, semiotic ground, directionality and representational relationship (from Zlatev et al., under review)

Vehicle	Bodily	Ground	Directionality	Representation
Ostensive gaze	Yes	-	Yes	No
Proximal point	Yes	Indexical (+ Symbolic)	Yes	No
Marker	No/Yes	Indexical	No/Yes	No/Yes
Picture	No	Iconic (+ Symbolic)	No	Yes
Replica	No	Iconic	No	Yes

The fact that children at this stage do not (yet) understand iconic representations, even when executed in the “bodily” modality (i.e. gestures), was shown by Namy, Campbell & Tomasello (2004), in an experiment where 18-month old children associated equally well iconic as well as arbitrary gestures with specific objects (small toys of a car, rabbit, hammer and spoon), while 26-month-old children performed much better with the iconic gestures than the arbitrary ones. This can be explained by assuming that at 18-months children do not understand gestures qualitatively differently from the previous stages (e.g. the “baby signs”), i.e. as *imitated schemas associated with a particular object or event*. It is rather the element of communicative intent (“give me the object that we associated with THIS action”) that is the novel element.

How can this “failure” in iconicity to make a difference in comprehension be reconciled with the testified use of iconic gestures in production of children of the same age? Zlatev (in press), for example, found a total of 72 gestures that were classified as iconic in 60 minutes of spontaneous interaction between caregivers and 6 children at approximately 18 months, or 1.2 iconic gestures per minute. The answer is above all in *the definition of gesture*, adopted from Andr n (2010), which requires either “explicitly other-oriented action, with visible communicative intentionality” or for the act to be used as an “explicit sign” (with expression standing for a referent), *but not necessarily both*.

In the case of iconic gestures in the study, thus, if the act was used with a marker of communicative intent, even (stylized) performances of practical actions (such as KISS and HIT) were included. In addition, enactive “symbolic play” representations, such as FEED (mother with an empty spoon) occurred at that early age. An additional factor contributing to the presence of iconic gestures in this stage is imitation: in the study they were found to be more often (on average 30%) directly imitated from the actions/gestures of caregivers than either deictic or emblematic (conventional) gestures.

Thus, *pace* Piaget (1962), as well as Zlatev & Andrén (2009), it is *not the understanding of representations* (“the symbolic/semiotic/sign function”) that constitutes the major difference compared to the previous stage, but rather the understanding of communicative intent. This understanding is achieved not so much intellectually as a higher-order intention, but as bodily markers accompanying acts of communication, signaling *that* one is communicating intentionally, very often for the benefit of the addressee. Understanding *what* is being communicated is signaled by semiotic vehicles that are (usually) performed with the body, allowing them to be readily imitated and “typified” (Andrén 2010). This is a major step in semiotic development, as well as in intersubjectivity, since it allows the further synergistic interaction between communicative intent and semiotic vehicles, paving the way to the insight that objects, actions and events have “names” (not necessarily verbal) that are commonly known, i.e. conventional, and thus eventually to language.

Stage 4: Communicative, conventional representations (20-30 months)

To give a rough estimate of the linguistic competence of the three groups of children in the object-choice study reviewed above (Zlatev et al., under review), we asked parents to fill in the forms of the *Swedish Early Communicative Development Inventory* (SECDI) (Berglund & Eriksson 2000), providing measures of the children’s receptive and productive vocabulary. It is characteristic that the median score in the most comprehensive measure (asking if the children produced any of 710 common lexical items) was 35 for the 18-month old children and 305 for those at 24 months, an increase of 900%. This was a clear reflection of the well-known phenomenon known as the *vocabulary spurt*, occurring for most children in the second half of the second year: “At first their rate of vocabulary growth is very slow, but one typically sees a “burst” or acceleration in the rate of vocabulary growth somewhere between 16-20 months” (Bates 2002: 15). What can explain this highly accelerated growth? While the idea of a “symbolic insight” was popular in earlier analyses of first-language learning (cf. Ingram 1989), more recent interpretations, including that of Bates (2002), have tended to downplay the phenomenon, and to attribute it to non-linear dynamics in rates of learning.

However, there are other indications that a cognitive-semiotic “reorganization” takes place between 18 and 26 months. Concerning the rather surprising results of Namy et al. (2004), summarized above, the

authors suggest the following explanation: “At 26 months, children have developed more rigid expectations than their younger counterparts about the forms that object labels may take” (ibid: 54). In other words, they propose that infants expect vocal labels not to sound like what they refer to, but that gestures, when used as labels, should resemble their referents. It can be noted that this explanation presupposes that during this stage, infants have some degree of explicit awareness (if they are going to have different expectations) *that words and gesture are used “as labels”, i.e. as signs*. The reason that iconic and arbitrary gestures were both associated with objects in the previous stage can actually be explained by assuming, as suggested earlier, that they were learned as associations, rather than as “explicit signs”, a possibility that Namy et al. do not consider. Still it would be consistent with their proposal of a re-organization in “symbol-learning” towards the end of the second year.

Furthermore, in our previous study of the development of children’s gestures from 18 to 27 months in three Swedish and three Thai children, which we analyzed in terms of the Mimesis Hierarchy (Zlatev & Andrén 2009), we also found evidence for a transition around 20 months: on average, this was the age when (i) deictic gestures, produced together with deictic expressions and nominals peaked, while (ii) what seemed like iconic gestures decreased, and (iii) emblematic (conventional) gestures suddenly increased. From then on, until 27 months, all these tendencies were reversed: the rates of deictic and emblematic gestures decreased, while iconic gestures (mostly cases of symbolic play), increased, along with measures of the children’s linguistic proficiency (vocabulary, MLU).

The explanation of this apparent reorganization that we offered was “a more or less *explicit understanding* (insight) that the meaning of the sign (gesture or word) is common to oneself and the addressee, i.e. the sign’s *conventionality*” (ibid: 384) which was qualified as “a kind of “symbolic insight”, not in the sense that the children did not use any signs prior to that, but that they grasped, at least partially, the nature of semiotic norms (conventions) around this time” (ibid: 396). Given the empirical findings mentioned earlier, and the theoretical advantages of distinguishing communicative intent and semiotic vehicles, as well as different levels of conventionality (Andrén 2010), this interpretation should be modified.

The present proposal is rather that the common denominator to the vocabulary spurt, the U-curve in interpreting arbitrary gestures, and the observed gestural reorganization, is most precisely captured by the original term *symbolic insight*, comprising the realization both that (i) “things have names”, and (ii) that these names are common, i.e. conventional, and thus at least to some degree normative. Otherwise, it is difficult to account why children at 26 months should have different expectations with respect to words and gestures: the first being typically “arbitrary”, while the latter typically “iconic”. This proposal has the further advantage of not requiring a “normative insight” at 20 months, which is indeed rather implausible, and a higher level of conventionality can build on the verbal and gestural

schemas acquired gradually through imitation/mimesis, since the onset of Stage 2. What makes the conventionality of “labels” more normative than that of actions, or mimetic schemas, is that misuse will tend to lead to misunderstandings, and frustrations of communication (“I want the DOG, not the BIRD”). Thus, semiotic normativity comes for free, so to speak, with the symbolic insight, though of course this is only its developmental onset. Throughout this stage, grammatical norms begin to be acquired, with piece-by-piece imitations, and “creative” generalizations (Tomasello 2003), the child proceeds to re-construct the linguistic system of the community throughout the duration of this stage, which should be viewed as a highly transitional stage, with upward borders that are somewhat diffuse.

Stage 5: Language-mediated folk-psychology (2.5 years -)

It is difficult to pinpoint the onset of truly “creative” language use, since even among adults utterances are often imitations and permutations of what they have already experienced. Still, it is clear that around the middle of the third year, children indeed say things that surprise caregivers. For example, at 3 years my son, after coming home from the first night-time car-trip in his life, commented: *Cars make the moon go*. It was not until sometime later that we realized that he was referring to the “apparent motion” of the moon behind the trees, when looking out of the window of a moving car.

Once children have developed a sufficiently expressive “conventional-normative semiotic system for communication and thought” (i.e. the definition of *language* adopted by Zlatev 2008b), this inevitably has repercussions for their understanding of “social cognition”, yielding an additional stage in the development of intersubjectivity. The following properties of language have been suggested to be instrumental for performance on so-called “theory-of-mind” tasks. First, structural features like mental predicates (verbs like *think, believe, know*) and sentential complement constructions (Astington & Jenkins 1999). Second, discursive features like disagreements, repairs and meta-linguistic discourse (Lohmann & Tomasello 2003). Third, Hutto (2008) has argued that linguistic proficiency brings first apprenticeship and then mastery in understanding and producing narratives,³ and it is through these that children, at least from their fourth year, begin to understand the folk-psychology of beliefs and desires, allowing them e.g. to pass “false-belief” tests. As Nelson (2003) has further argued, knowledge of “cultural myths and social narratives” has a constitutive role for forming autobiographical memories. This observation highlights an important theoretical point: that subjectivity and intersubjectivity are co-dependent categories, and that development in one is intertwined with development in the other. Thus, the stage-model here presented can also be regarded

³ This focus on language-mediated narratives is most in line with Donald’s original proposal to call language-dominated culture and cognition “mythic”.

as a model of the development of self-hood, which explains why it tallies to some extent with the one offered by Stern (1998).

4. Conclusions

In this article, I have elaborated on, and corrected some interpretations from previous work on the relationship between bodily mimesis and intersubjectivity (Zlatev 2008a) and the application of the Mimesis Hierarchy model to semiotic development in children (Zlatev & Andr n 2009). The model of (at least) five more or less distinct stages stands in contrast to those who treat the development of intersubjectivity as gradual, with most capacities essentially present “from the start” and only in need of unfolding (Trevarthen 1979). Or alternatively, as a two-stage process, the first stage a matter of enactive perception and interaction, and the second - introducing narrative (Gallagher 2005; Hutto 2008). It is, of course, also quite distinct from those operating with the concept of “theory of mind”, either of the theory-theory or simulation-theory variety (cf. Zlatev et al. 2008). By being a multi-level model, it is most similar to that of Stern (1998) on the development of “the sense of self”, and to Tomasello’s (1999) model of the “cultural origins of human cognition”, including both pre-linguistic and linguistic factors, as well as to Nelson’s (1996) application of Donald’s evolutionary model to development. Though, naturally, it differs in most of the specifics.

As in previous treatments, the MH-model focuses on bodily mimesis, its “precursors” (empathetic perception) and “post-developments” (conventionality, language and narrative). Mimesis is pivotal, as in Donald’s evolutionary model, since it provides the basis for the development of (i) conventions (through imitation), (ii) intentional communication, and (iii) for bringing the two together in communicative, shared representations (signs). What Donald states for evolution, applies equally well for development: “Language is different from mimesis, but it has mimetic roots. It is a collective product and must have evolved as a group adaptation, in the context of mimetic expressive culture. Given the conventional, collective nature of language, it could not have emerged in any other way.” (Donald 2001: 274)

The main difference from the previous applications of these ideas has been in the treatment of the concepts of *representation* and *communicative intent*. Under the influence of Piaget (1962), I previously regarded representations as emerging from the imitation of practical acts, i.e. as properties of dyadic mimesis (Stage 2), and treated children’s first gestures as “externalizations” of these, with communicative intentions “added on” to yield triadic mimesis (Stage 3). With the onset of semiotic normativity (Stage 4), communicative focus turns to language, and gestural signs undergo a consequent reorganization. Due to the empirical findings reviewed in Section 3, and a less “mentalist” and more bodily-enactive and social-oriented perspective, I have here proposed a more or less reversed

sequence: Stage 2 gives rise to imitation and *mimetic schemas* (Zlatev 2007, in press), but the first gestures (and vocalizations) of children are neither externalizations of these “internal representations”, nor fully-fledged representations/signs on their own, but action schemas bi-directionally *associated* with particular contexts. That is why the onset of intentional communication occurs in Stage 3 with *pointing and other deictic gestures (such as showing), which are not representations or fully-fledged (explicit) signs*, but rather performative communicative acts, accompanied with makers of communicative intent. It is first in Stage 4 that the proto-representations of Stage 2 (Piaget’s internalized imitation-based “symbols”, my mimetic schemas) and the communicative intent of Stage 3 are combined to give rise to communicative iconic gestures, and more generally to the “insight” of using communicative, shared representations, or what is variously referred to as “symbols” (Tomasello 1999; Namy et al. 2004) or “signs” (Zlatev 2009). While this account would undoubtedly undergo further modifications, it resolves a number of difficulties inherent in the previous one.

Finally, I hope to have demonstrated that models of semiotic development need not be focused on specific skills and time periods, but can follow in the tradition of Piaget and propose more general, integrational accounts, which can lead to scientific progress, despite the risk of being wrong in many of the particulars.

Acknowledgements

I wish to thank my colleagues at the Centre of Cognitive Semiotics (CCS) at Lund University, for the ideas, discussions, and collaborations in empirical studies serving as the basis for this article, and in particular (in alphabetical order), to Mats Andrén, Ingar Brinck, Sara Lenninger, Elaine Madsen, Joel Parthemore, Tomas Persson, Michael Ranta, Susan Sayehli and Göran Sonesson. In addition, I wish to acknowledge my long-term collaborators on the subject of intersubjectivity, and co-editors of *The Shared Mind*, Esa Itkonen, Tim Racine and Chris Sinha (despite our many disagreements). And last, but not least I thank Christiane Moro and Paul Bouissac for the positive interactions that provoked the publication of two different versions of this article, in French and English.

References

- Acredolo, L. & S. Goodwyn (2000). *Baby Minds: Brain-building Games your Baby Will Love*. New York: Bantam.
- Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development*, 59, 450–466.
- Andrén, M. (2010). *Children's Gestures between 18 and 30 Months*. Published Doctoral Dissertation. Department of Linguistics, Lund University.

- 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.
- Baldwin, J. M. (1894) Imitation: A chapter in the natural history of consciousness. *Mind*, 3, 26-55.
- Bard K. A., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A. and Matsuzawa, T. (2005). Group differences in the mutual gaze of chimpanzees. *Developmental Psychology*, 41, 616-624.
- Bates, E. (2002). On the nature and nurture of language. Unpublished report. UCSD.
- Bates, E. Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The Emergence of Symbols. Cognition and Communication in Infancy*. New York: Academic Press.
- Berglund, E., & Eriksson, M. (2000). Communicative development in Swedish children 16 - 28 months old. The Swedish early communicative development inventory - words and sentences. *Scandinavian Journal of Psychology*, 41(2), 133-144.
- Brinck, I. (2003). The pragmatics of imperative and declarative pointing. *Cognitive Science Quarterly*, 3(4), 429-446.
- Carpenter, M., Nagell, K. & Tomasello, M. (1998). *Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age*. Monographs of the Society for Research in Child Development, Vol. 63, No. 4. Blackwell Publishing.
- de Waal, F.B.M. (2007). The 'Russian doll' model of empathy and imitation. In S. Bråten (ed.), *On Being Moved: From Mirror Neurons to Empathy*, 35-48. Amsterdam: Benjamins.
- DeLoache, J., S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences*, 8(2), 66-70.
- 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.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.
- Garnder, H. (1992). *Multiple Intelligences*. New York: Basic Books.
- Grice, P. (1957). Meaning. *Philosophical Review*, 66, 377-88.
- Iacoboni, M. (2008). *Mirroring People: The New Science of How we Connect with Others*. New York: Faber, Straus & Giroux.
- Ingram, D. (1989). *First Language Acquisition: Method, Description and Explanation*. Cambridge: Cambridge University Press.
- Iverson, J.M. and Goldin-Meadow, S. (1998). *The Nature and Functions of Gesture in Children's Communication*. San Fransisco: Jossey-Bass Publishers.
- Lenninger, S. (2012). *When Similarity Qualifies as a Sign: A Study in Picture Understanding and Semiotic Development in Young Children*. Published Doctoral Dissertation. Department of Semiotics, Lund University.
- Lohmann, H. & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development*, 74/4, 1130-1144.

- Mandler, J. (2004). *The Foundations of Mind. Origins of Conceptual Thought*. Oxford: Oxford University Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: University of Chicago Press.
- Meltzoff, A. & Moore, M. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75-78.
- Meltzoff, A. & Moore, M. (1983). Newborn infants imitate adult facial abilities. *Child Development*, 54 (3), 702-9.
- Merleau-Ponty, M. (1962). *Phenomenology of Perception*. London: Routledge.
- Moore, R. (under review). Cognizing communicative intent. *Mind and Language*
- Moro, C. (2011). Material culture, semiotics and early childhood development. In M. Kontopodis, C. Wulf & B. Fichtner (Eds.), *Children, Development and Education: Cultural, Historical, Anthropological Perspectives*, 57-70. London, New-York: Springer Verlag.
- Namy, L., Campbell, A. L. & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, 5(1), 37-57.
- Nelson, K. (1996). *Language in Cognitive Development. The Emergence of the Mediated Mind*. Cambridge: Cambridge University Press.
- Nelson, K. (2003). Self and social functions: Individual autobiographical memory and collective narrative. *Memory*, 11 (2), 125-136.
- Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books.
- Piaget, J. (1962). *Play, Dreams and Imitation in Children*. Translated from *La formation du symbole chez l'enfant*, Delachaux et Niestlé, 1945. New York: Norton.
- Preston S.D. & de Waal F.B.M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25, 1-72.
- Reddy, V. (2003). On being an object of attention: Implications of self-other-consciousness. *Trends in Cognitive Science*, 7/9, 397-402.
- Reddy, V. (2005). Before the “third element”: Understanding attention to self. In *Joint Attention*, N. Klein (Ed.), 85-109. Oxford: Oxford University Press.
- Rodriguez, C. & Moro, M. (2008). Coming to agreement: Object use by infants and adults. In J. Zlatev, T. Racine, C. Sinha and E. Itkonen. (Eds.) In *The Shared Mind: Perspectives on Intersubjectivity*, 89-114. Amsterdam: Benjamins.
- Siegler, R. (1996). *Emerging Minds*. Oxford: Oxford University Press.
- Sperber, D. & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd edition). Oxford: Blackwell.
- Stern, D. N. (2000). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. New York: Basic Books.

- Tomasello, M. (1995). On the interpersonal origins of the self-concept. In U. Neisser (Ed.) *The Perceived Self. Ecological and Interpersonal Sources of Self-knowledge*. Cambridge: 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. Call, J. & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, 68(6), 1067-1080.
- Trevarthen, C. & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action, Gesture and Symbol: The Emergence of Language*, 183- 229. London: Academic Press.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech*, 321-347. Cambridge: Cambridge University Press
- Zlatev, J. (2007). Intersubjectivity, mimetic schemas and the emergence of language, *Intellectica*, 2-3 (46-47), 123-152.
- Zlatev, J. (2008a). The coevolution of intersubjectivity and bodily mimesis. In J. Zlatev, T. Racine, C. Sinha and E. Itkonen. (Eds.) In *The Shared Mind: Perspectives on Intersubjectivity*, 215-244. Amsterdam: Benjamins.
- Zlatev, J. (2008b). From proto-mimesis to language: Evidence from primatology and social neuroscience. *Journal of Physiology – Paris*, 102, 137-152.
- Zlatev, J. (2009). The Semiotic Hierarchy: Life, Consciousness, Signs and Language, *Cognitive Semiotics*, 4, 169-200.
- Zlatev, J. (in press). Image schemas, mimetic schemas and child development. *Gesture*, Special issue devoted to image schemas.
- Zlatev, J. and M. Andrén (2009). Stages and transitions in children's semiotic development. In J. Zlatev, M. Andrén, C. Lundmark and M. Johansson Falck (Eds.) *Studies in Language and Cognition*, 380-401. Newcastle: Cambridge Scholars.