Mimesis:
The “missing link” between signals and symbols in phylogeny and ontogeny?

Jordan Zlatev
Department of Linguistics
Lund University, Sweden
Jordan.Zlatev@ling.lu.se

1. Introduction

The purpose of this chapter is to explore the idea expressed in the title: that mimesis, the voluntary use of one’s body as a representational device, played a crucial role in the origins of human language and thought, and continues to do so, though possibly to a lesser extent, in their development in children.

The basic idea is not new, but has been recently elaborated and defended with respect to human phylogeny by Donald (1991, 2001), and applied to child development by Nelson (1996), and (independently) by Zlatev (2001a, 2001b). At the same time, with the strongly rekindled interest in the evolution of language, similar but competing evolutionary theories have instead attributed a pivotal role to praxis (Corballis 1991), gesture (Armstrong, Stokoe and Wilcox 1995), imitation (Tomasetello 1999) and ritual (Deacon 1997). I would suggest that the battle is at least in part terminological, reflecting the importance of “branding” in modern capitalist culture (Klein 2000), with theorists keen on emphasizing their own favorite concept at the expense of those of competitors. My choice falls on Donald’s work, since it stands out as the most integrative, and his concept of mimesis as most comprehensive. It has been defined as follows:

Mimetic skills or mimesis rests on the ability to produce conscious, self-initiated, representational acts that are intentional but not linguistic. (Donald 1991: 168)

Mimesis is the result of evolving better conscious control over action. In its purest form, it is epitomized by four uniquely human abilities: mime, imitation, skill, and gesture. (Donald 2001: 263)

In the latest work, these four mimetic abilities are placed in a tentative evolutionary hierarchy. Mime is considered simplest since it (only) involves a reenactment of an external event. Imitation is more complex because it requires grasping the purpose
(intention) behind the modeled event, as also emphasized by Tomasello (1999). **Skill**, is closely based on the previous two, since it involves the rehearsal (self-imitation) of a series of hierarchically organized mimetic acts, until some ideal standard is reached. Finally, **gestures** are explicitly communicative, often elaborately sequenced body movements, requiring a shared understanding of their meaning for communication to succeed.

This latter definition of gesture is considerably narrower than that of Armstrong et al. (1995), who define it as “a functional unit, an equivalence call of coordinated movements that achieve some end” (ibid: 46). This includes just about anything from a goal-directed leap to the “visual gestures” of Sign\(^1\) languages, and the “invisible vocal gestures” of spoken languages, and therefore represents a concept so general as to be practically vacuous. However, it does serve the gradualist theory of language origins espoused by Armstrong et al. (1995), according to whom the differences between “ape gestures” and Sign language (or vocal language, for that part) seem to be a matter of greater complexity and conventionalization. If this were indeed the case, then there is simply no gap for any “missing link” to fill.

But there are some compelling reasons to believe that there is one. Donald clearly distinguishes the **event representations** of apes from the language-based representations of human culture. So does Deacon (1997) referring to the cognitive and communicative categories of animals (mistakenly) with the Peircean terms **“icon”** and **“index”** (Pierce 1931-35). Likewise, Tomasello (1999) distinguishes between “gestural signaling” and gestures proper, arguing that apes (in the wild) are only capable of the first since they lack an understanding of others as psychological beings, which constitutes a requirement for intentional communication (Grice 1957). Indeed, the differences between animal **signals** and human **symbols** (Vygotsky 1978, Deacon 1997, Tomasello 1999, Sinha 2001, this volume), some of which are summarized in Table 1, are so considerable that it is almost impossible to envision a single transition from the first to the second. This has provided fuel for otherwise evolutionary anomalous, strongly nativist, suggestions of a sudden grand mutation, creating language almost **ex nihilo** (Bickerton 1990, Piattelli-Palmarini 1989). It is also characteristic that Deacon’s attempt to explain the transition to symbolic reference with the onset of marriage ritual is where his theory is least convincing (cf. Armstrong 1998).

---

\(^1\) I adopt the convention of spelling “Sign” with a capital letter, when it used to refer to the manual-brachial language of the deaf.
Table 1. Some differences between signals and symbols in terms of relevant cognitive characteristics, making the two categorically distinct (cf. Vygotsky 1978, Deacon 1997, Sinha 2001, this volume). Note that there can be quantitative differences within each category.

<table>
<thead>
<tr>
<th></th>
<th>Signals</th>
<th>Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examples:</td>
<td>bee “dance”, bird song, vervet monkey calls</td>
<td>Kanzi’s “lexigrams”, words, mathematical symbols</td>
</tr>
<tr>
<td>Degree of learning</td>
<td>None (genetically determined) or limited</td>
<td>Extensive</td>
</tr>
<tr>
<td>Conscious control over production</td>
<td>None, or highly limited</td>
<td>High (in normal cases)</td>
</tr>
<tr>
<td>Interpretation</td>
<td>(Relatively) fixed response</td>
<td>Flexible, “negotiable”</td>
</tr>
<tr>
<td>Contextuality</td>
<td>Tied to a particular context (stimulus setting)</td>
<td>Flexible, relatively independent from specific context</td>
</tr>
<tr>
<td>Systematicity</td>
<td>None, or very limited</td>
<td>High</td>
</tr>
</tbody>
</table>

A chief advantage of Donald’s theory is that it has suggested how the bridge between animal and human communication and cognition can be built not by a miraculous “chance mutation”, but by mimesis which gave rise to the first public representations and thus serving as the basis for a culture of shared scripts, ritual, pedagogy and gestural communication. On their part these cultural skills prepared the transition to the oral-linguistic culture of *Homo sapiens*.

It is this hypothetical role of a “missing-link” played by mimesis that will be the topic of this chapter. In Section 2, I will review some of the evidence for a “mimetic stage” in evolution presented by Donald, but also by others. Since ontogenetic development can be expected to resemble in broad lines – though not to “recapitulate” – phylogeny due to its epigenetic character, we should expect mimesis to be strongly present there too, and Section 3 will discuss some relevant phenomena in child development. While these two sections deal with the hypothesis empirically, Section 4 will briefly address the question from a more conceptual perspective: given the consensual, intersubjective nature of language, and the nature of our bodies and brains, could mimesis even be a necessary precursor to language? Finally, Section 5 summarizes the conclusions of the study, and points out implications.

Before proceeding, I wish to call attention to a typographical convention that will be adhered to: Since the text is full of many theoretically loaded terms, some of which have different everyday meanings, for the sake of clarity I will continue to place such terms in **boldface**, and whenever possible will add a definition. When a definition is absent the reader may regard the boldface as a “warning sign”, and if unsure to consult the sources.
2. Evidence for a “mimetic stage” in hominid evolution

2.1. Two major transitions in biological and cultural evolution

There is considerable agreement among paleontologists and physical anthropologists that with *Homo habilis/erectus* some 2 million years BP (before present) there was both a cultural and a neuroanatomical “jump” in hominid evolution. The transition had begun earlier, with the branching off from the chimpanzee line about 5-6 million BP, and gained momentum with the different species of the genus *Australopithecus* (who were all bipedal) some 4,000,000 years BP. However, the biology and culture of *Homo erectus* was the first that showed a qualitative break with the animal world. Their brain/body ratio was twice that of apes – though still 70% that of modern humans. They constructed fairly elaborate tools, migrated out of Africa and colonized most of the world, using fire and base camps. At the same time, their vocal tracts were still similar to those of apes, and there is no evidence that they possessed any kinds of vocal language. They did have a homologue to *Broca’s area* – implicated in the use of language in modern humans – in the left hemisphere, but this has also been shown to be involved in *praxis* (Corballis 1991) and *imitation* (Rizzolotti and Arbib 1998), as explained below. As summarized by Donald (2001: 267):

> For over two million years hominids shared food, disseminated skills by imitation, created mini-industries for activities such as tool manufacture, and held group behavior to very stable long-term patterns, such as the continuous habituation of a site and the maintenance of fire. These things are not easy to achieve, and there is no good evolutionary rationale for attributing them to language.

But this long period of relatively continuous development was interrupted by a second transition some 200,000 - 150,000 years ago with the advent of *Homo sapiens*: brain size, especially due to *prefrontal cortex* expansion, radically increased and reached modern proportions and the vocal tract attained its present form. This coincided with a cultural explosion and a new exodus “out of Africa” leading to a colonization of the world that was to be total, wiping out any predecessors such as the Neanderthals in the process. As many have argued (e.g. Lieberman 1991), there are good reasons to think that a crucial advantage in *Homo sapiens* compared to other species and predecessors was the possession of a high-speed, articulate vocal language. But in that case, what was it that served the basis for *Homo erectus* culture? Mimesis, rather than either spoken language or Sign language, stands out as a likely candidate.

2.2. “Mimetic vestiges” in modern cultures and brains

A second sort of evidence appealed to by Donald is based on the metaphor of “vestiges” – traces of earlier stages of evolution in modern human culture and communicative behavior. On an intuitive level, our fascination with many forms of art (dance, theatre, pantomime), games (“follow the leader”, charades) and athletics testifies to the centrality
of mimesis in our daily life – in the form of mime, imitation, motor skill, or combinations of these.

But perhaps the clearest evidence for a mimetic “substratum” underlying human verbal communication is the ubiquity of gesture, especially of the type called by Kednon (1988) gesticulation and by McNeill (1992) speech-associated gestures (cf. Gullberg 1998). Such gestures are usually classified as iconic and metaphoric (which depict concrete or abstract concepts respectively) or rhythmic (“batons” or “beats”). What is curious is that such gesturing is largely performed unconsciously, and at least in part unintentionally: even blind people gesture when speaking. The fact that it is closely synchronized with speech has lead McNeil to claim that the two communication systems share a common underlying representation, and therefore that gesture is no less “verbal” than speech: “Gesticulation and language are seen as equipotent reflections of thought with different output channels. They are interdependent and one is not a translation of the other”, as summarized by Gullberg (1998: 62). This interdependence has been further supported by recent evidence for gesture differences between Spanish and English speakers which correlate with semantic characteristics of the respective language (McNeill 2000).

On the other hand, McNeil (1992) himself states that speech-associated gestures lack hierarchical structure and semantic compositionality – two characteristics usually considered defining of linguistic representation – which places in doubt the “equipotentiality” of language and gesture. Evidence from aphasia shows that while loss of speech is often accompanied by lack of appropriate gesturing, this is not so in all cases, in particular in individuals with global aphasia, completely unable to produce and understand speech (even as an internal monologue), but maintaining the ability to communicate with iconic gestures of the type that are usually “speech-associated”, as well as through miming, pointing and by using standardized, culture-specific emblems. All this supports the thesis “that mimetic representation is the more fundamental level of representation” (Donald 1991: 224), separate from, but in normal conditions interacting with, and under the influence of the evolutionarily more modern linguistic system.

A fairly clear demonstration of the fundamental though “sub-verbal” character of mimetic representation can be seen in the gestural communication of deaf people who are not exposed to Sign. Such individuals, e.g. deaf children of speaking parents, interact with their hearing environments or with other “Signless” deaf people through a form of pantomime that has come to be called homesign, studied extensively by Goldin-Meadow and her colleagues. It has been established that over time, processes of simplification and standardization make the communication system more arbitrary and compositional, i.e. less mimetic and more linguistic (Morford, Singleton, Goldin-Meadow 1995):

After a language has undergone several generations of change, very little attention is paid to the iconicity of symbols by the language user at all. Thus, it is only when systematicity is given priority over iconicity that symbols are transformed into language. (ibid: 330)

---

2 Such as the “OK” sign performed by touching the tops of index finger and thumb.
On the one hand, this can be said to show the difference between mimesis and language, with the transition between the two providing a kind of “recapitulation” of evolution. On the other hand, the readiness with which speech-impaired children acquire Sign language, or even “create” one in the absence of a model, has lead some researchers to propose that not only mimesis, but a full-blown sign language preceded spoken language in evolution (Armstrong et al. 1995, Stokoe 2001). This is possible, but by no means necessary. If we assume that mimesis and speech constitute interacting systems (as the evidence suggests) we can see how a brain adapted for the acquisition of symbols represented in speech, such as the brain of Homo sapiens, would also be one adapted for the acquisition of Sign through the “feedback” projections of the more modern system (language) over the more ancient one (mimesis), thus recruiting it for symbolic communication in the context of deafness. This is also consistent with the evidence outlined in Section 2.1 that Homo erectus hardly had access to a full language, either spoken or Sign.

2.3. Manual control and bipedalism

One could argue that the evidence reviewed above is consistent with the possibility that vocal language developed in parallel with and independently from mimesis and gesture, linking with it only after prolonged evolution (and never completely, as the data from aphasia suggest). One such scenario is suggested by Deacon (1997) according to whom the first symbols arose through pair-boding ritual as early as the genus Australopithecus and started a chain of brain-culture co-evolution – all this with little or no role for mimesis. This scenario is also consistent with evidence for a gradual evolution of the vocal tract from Homo erectus to anatomically modern Homo sapiens (Lieberman 1991).

However, one difficulty in assuming that the first “symbols” were vocal is that in that case they would have originated from the primate call system, which is far less subject to conscious control than the limbs. This is illustrated by the anecdote told by Deacon (1997) of a chimpanzee who has found food in a secluded place and attempts to muffle his involuntarily produced food-call with its hand, in order to be able to enjoy the food alone. Recent neurobiological evidence shows that it is a region in cingulate cortex that controls primate calls and that this region is not at all homologous (i.e. having the same origin) to Broca’s area in modern humans, which as well-known (though not uncontested) plays a key role for the grammatical organization of language. Broca’s area has rather been argued to be homologous to areas in the pre-motor cortex involved in manual-brachial (hand-arm) control and action recognition (Rizzolatti and Arbib 1998).

With the onset of bipedalism in Australopithecus the degree of manual control and proficiency would have increased even more. While there are many other possible evolutionary pressures to explain the transition to bipedalism – e.g. change of ecological niche, childrearing and tool-making – mimetic gestures could clearly have been a factor for establishing the transition. It is quite possible to imagine a scenario in which an existing manual communication system eventually “recruited” vocalization, bringing it under voluntary control and providing the selection pressures for lowering the larynx. The advantages for this latter change must have been quite strong to counterbalance the
negative effects of more likely choking, which would only have been the case if there had been language-like communication prior to anatomically modern Homo sapiens, which is strong evidence against a relatively late emergence of all linguistic skills (Johansson 2002).

2.4. Co-lateralization for language and “praxis”

As well-known, language skills are “lateralized” in about 90% of all people, meaning that most of the crucial (though not all!) areas for language production and comprehension are located in the left hemisphere, above all the classical two: Broca’s area in the left frontal pre-motor cortex, and Wernike’s area in the left temporal gyrus. A similar proportion of the population is “right-handed”, in the sense that precise and complex motor activities are better performed with the right hand (and foot), both of which are under the control of the left hemisphere. Among right-handed people, the co-relation with left-hemisphere specialization for language is nearly perfect. With left-handers, the majority still have left-hemisphere dominance for language, but there is a sizable minority who don’t. What is the reason for this strong, if not perfect correlation between language dominance and handedness? Corballis (1991) has proposed that not just language, but all behavioral systems involved in assembling complex sequences of actions (including visual routines) are controlled by the left hemisphere, which is endowed with a specific cognitive adaptation, termed generative praxis. The claimed origin of this capacity is the refinement of manual skill, and especially tool-making in Homo erectus some 2 million years BP. Greenfield (1991) is even more specific in claiming that Broca’s area itself is implicated as the cerebral cortical basis for the hierarchical organization of language and manual object combination, presenting evidence that the two abilities develop simultaneously in human ontogeny.

All this is, of course, consistent with a central role for mimesis, since motor skill is according to Donald a mimetic ability. But while mimesis in general implies the capacity for representation, praxis does not. And since representation is a central characteristic of language, it remains unclear why language should “piggyback” on, or develop from structures for praxic skill. What is necessary is to show that the “proto-Broca” areas were involved not just in the control of action sequencing, but in semiotic activity, i.e. the use of signs – in communication, thought or both.

The missing piece of the puzzle could possibly be provided by the recent discovery of the “mirror-neuron” area F5 in the pre-motor cortex of the monkey brain, which mediates between observed actions and performed action, and which has been persuasively argued by to be homologous with Broca’s area in modern humans, as suggested below.

2.5. “Mirror neurons”, iconicity and parity

Rizzolatti and Arbib (1998) and Arbib (in press) sketch the following scenario for the role of the mirror neuron system in the “proto-Broca” area in the development of
mimesis. Initially it was used for action recognition, as it still is in monkeys. In the great apes, and in our common ancestor some 6 million years BP, it was extended for the purpose of imitation of goal-directed actions. A crucial step was realized in hominids, who began using the system for the production, perception and imitation of actions without the involved objects, in other words for pantomime. This according to Arbib (in press), would have been the first form of explicit representation. For example if hominid A performs a grasping motion towards an imaginary fruit, hominid B would recognize his actions and would be likely to understand it as an expression of a certain intention (provided at least a rudimentary “theory of mind”). Notice that this scenario is different from Donald’s (2001) mimetic hierarchy (mime > imitation > skill > gesture), by placing (panto)mime, which is explicitly representational, after imitation/skill, which are not. This reordering is quite plausible from a Vygotskian “Outside-Inside” perspective, according to which “higher mental functioning” first emerges between people in communication, before it can mediate between the subject and the physical environment, and even later within the individual’s own mind.

An advantage of the above hypothesis is that it does not require a major “insight” that the body could be used as a representational tool from some imaginary early hominid for mimesis to get off the ground. Since learning by imitation, and a social life dependent on it, would have already been quite established, at the latest by the time of Homo erectus, all that would have been necessary is for a “teacher” occasionally to perform a particular action without the necessary object, to make this a highly iconic sign of that action. Since the mirror system provides a basis for parity – an externally observed and a self-produced action are in some sense categorized as equivalent – the sign, as well as the full action it signifies will be “shared” between sender and receiver.

Another advantage is that in this way we can see how the “germ” of grammar would have been inherent in the first communicative acts, which could have consisted of (eventually stylized) actions with “slots” for the participant roles: the Actor, the Patient, the Instrument etc. With standardization, these different parts would have been further differentiated from the action itself, which is conceptually much simpler than imagining how “things”, “actions” and “properties” could be segmented and named individually and combined afterwards. Armstrong et al. (1995) point this out as an argument for the gestural origin of language:

Not only visible gestures … often resemble what they signify, they are also likely to have originated … [through the discovery] that certain actions observed were, and others could be, mimicked with manual-brachial actions. … [N]oting so far seems to have as much potential as visible actions involving an actor and its action for guiding hominids to the act of symbolizing relations as well as things. (ibid: 22-23)

The hypothesis is thus that grammatical structure was initially highly iconical, and this rhymes with the observation that it largely still is so in Sign languages (Armstrong et al. 1995) and to a lesser extent even in spoken languages (Haiman 1985, Itkonen, this volume). Degree of iconicity, however, should not be taken as a simple indication of how
ancient a communication system – and thus that full Sign language intervened between mimesis and spoken language.

2.6. Summary and unresolved issues

Even the cursory review of some of the evidence for the role of mimesis in human evolution presented above shows that mimesis has many of the properties necessary for bridging the gap between animal signal systems and human symbolic communication, and particularly language:

- (partial) generativity
- intentional communication
- public representation
- parity
- iconicity

On the one hand it is readily imaginable how the early hominids developed such a system, and on the other, how it was further transformed towards greater **symbolicity** (cf. Section 1). While it is not impossible that this later development went through a “stage” of more or less modern-like Sign language, it was pointed out that this inference is by no means necessary, and probably less likely than a scenario in which speech began to be “recruited” for disambiguating between gestures fairly early, and with time and process of **“Baldwinian evolution”** to take over as the main vehicle of communication, interacting with but separate from mimesis. If this was so, then the transition from gestural to vocal communication was not only one of modality change, but one of **qualitative** semiotic change, where a mimetic-iconic system is supplemented by a symbolic one. In this case the motivation for the transition would not only have been “external”: e.g. freeing the hands and visual attention for other means, communicating over barriers and in the dark, better **pedagogy** in tool-making (cf. Givon 1998, Corballis 2002), but also “internal”, since the vocal medium would automatically have introduced greater **arbitrariness** (Saussure 1916, Morford et al. 1995) in the mapping between expression and meaning and greater **differentiation**, thus pushing the development of “protolanguage” to higher symbolicity.

The discussion on whether mimesis first developed into a manual-visual language, or whether it served as a stepping stone to oral language directly parallels the disagreement on whether gesture and language constitute a single system (McNeill 1992), or rather are different evolutionary layers of the **“hybrid mind”** (Donald 1991). If gesticulation and language do differ in terms of hierarchical structure and compositionality, as McNeill himself claims, and if the evolution of oral language transformed mimesis into something qualitatively different: a conventional symbolic system, then despite claims to the contrary, gesticulation is more mimetic and less “verbal” than speech (and Sign).

On the other hand, the major drawback of the concept of mimesis, as used by Donald is that it is too inclusive. As shown there is quite a difference between skill, imitation, mime and gesture (and even more so, when e.g. ritual and dance are included) to regard it as a
single “stage” in evolution. More research is necessary in order to determine the most likely progression, and some of the relevant evidence may come for child development, to which we turn in the next section.

3. Evidence for mimesis in ontogenetic development

In contrast to both nativist and empiricist approaches to human cognitive development, the paradigm of epigenesis (Piaget 1971, Waddington 1975, Sinha 1988, Zlatev 1997) proposes that ontogeny can be explained as an ordered succession of developmental stages toward higher complexity that emerges through interaction with the environment, influenced by but not determined by the genes. It is a strong alternative since it is scientifically parsimonious (e.g. no commitment to an innate Universal Grammar, cf. Haukioja this volume), evolutionary plausible (Müller 1996), descriptively adequate (Nelson 1996), and theoretically attractive in providing a way to integrate the “natural” and “cultural” lines of human development (Vygotsky 1978, Tomasello 1999).

Another advantage, less often noticed, is that epigenesis also offers a basis for integrating studies of evolution and ontogeny, since it provides a long-missing rationale for Haeckel’s (1874) famous “biological law of recapitulation”, albeit in a weaker form: Since evolutionary more ancient cognitive structures (in general) require less experience in order to mature than evolutionary more modern structures, the early stages of ontogenesis will (in general) be dominated by evolutionary more ancient mechanisms, and the later stages by more modern mechanisms. In some cases the “lower” stages may even serve as an epigenetic prerequisite, i.e. a necessary step, for reaching the “higher” functions.

The question to ask in the present context is if mimesis – in its different manifestations – may be said to function as a “stage” in human ontogenetic development, and if it plays the role of an epigenetic prerequisite for the acquisition of symbols in general, and language in particular. Such a mediating role of mimesis in human development has been proposed independently by Nelson (1996) and Zlatev (2001a, 2001b), but with rather strikingly different conclusions: While according to Nelson mimetic representation dominates human cognition in early childhood, approximately from 1.5 to 4 years, in my application of Donald’s evolutionary model to ontogeny, I discerned an analogous “mimetic stage” ranging between the landmark developments of the “9 month revolution” (Tomasello 1999, cf. below) and the vocabulary and grammar spurts towards the end of the second year. This discrepancy calls for an explanation. After a review of some of the relevant evidence for the nature and timing of mimetic skills during the first years of life, I will try to determine which (if any) of the proposals is most consistent with the facts, and with the concept of mimesis itself.

3.1. Innate predispositions for imitation

Perhaps the most widely publicized results on infant imitation during the last 25 years have been those of Metzoff and Moore (1977, 1983, 1995) which demonstrated that newborn babies are capable of imitating simple movements involving mouth-opening,
tongue-protrusion and lip-protrusion, thereby showing that the human capacity to imitate is at least partially innate. The capacity to perform such “blind imitation” in the absence of a unitary sensory metric in which to compare one’s own action with that of another (which, notice, exists for oral-vocal imitation) is really quite amazing, and it is unsurprising that it is still sometimes contested. In terms of the subsequent findings of a “mirror neuron system” mentioned in Section 2, this inborn capacity can possibly be understood as a partial “pre-wiring” resulting in a mapping between the neural representations of the baby’s own body movements and visually induced representations of the movements of another human being, though how this can be realized in specific terms is still anybody’s guess. Granting the validity of the phenomenon, would it be justified to claim that mimesis is present from birth? Hardly, unless one is also prepared to describe the neonate’s imitations as “volitional acts of representation”, which even Meltzoff and Moore refrain from, given these movements’ limited and stereotypical character.

But in a more recent study Meltzoff and Moore (1995) report imitation of effortful, not naturally occurring behavior such as moving the tongue from side to side from as early as 6 weeks, and argue that infant imitation has the following features that distinguish it from mimicry: intentionality (ability to distinguish the goal state from the actual performance and correct the latter in terms of the former), selectivity (imitating different aspects of the model), creativity (original ways of reaching the same goal) and volitional control (ability to differ imitation, possibly indefinitely). This is quite a tall order, and if granted would imply the existence of mimetic skill soon after birth. But while the reported data strongly suggest intentionality (in the sense of “goal-directedness”) and some creativity, the other features are quite probably attributed due to an exceptionally “rich interpretation”. For true volition the 6-week old infant must possess self-consciousness, and there is no evidence to support this. Furthermore, mimesis possesses one more crucial feature, which would naturally follow with the above four, namely representation, the ability to distinguish between “signifier” (one’s own body) and “signified” (another’s action or some other event or object), and to realize that the first stands for the latter – a capacity that is even less imaginable in the new-born, as also admitted by Meltzoff and Gopnik (1993).

The most likely conclusion, consistent with the epigenetic standpoint, is therefore that Meltzoff and Moore’s studies of infant imitation have uncovered an adaptation for mimesis, but that these innate skills are only “bootstraps” – in the manner of sucking reflexes but on a higher level of complexity – which require adequate social interaction in order to lead to the development of imitation proper, possibly along with self-consciousness and reference, as suggested below.

3.2. Imitation, intersubjectivity and representation

As emphasized by Donald and Tomasello, true imitation can be distinguished from behaviors such as mimicking and stimulus enhancement by the fact that it involves the modeling of goal-directed behavior, which requires at least some understanding of the
model as an intentional agent. Tomasello presents much evidence that such an understanding develops first around 9 months, summarizing the transition with the memorable phrase “9-month revolution” and describing it as follows:

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. All of these behaviors indicate a kind of social-cognitive revolution: At 9 months of age infants begin to understand that other people perceive the world and have intentions and feelings towards it; they begin to understand them as intentional agents. (Tomasello 1995: 175)

On the other hand, the well-known infant developmentalist Threvarthen (1992) has argued persuasively for a kind of primordial intersubjectivity (sharing of experience) between infant and caregiver present almost from birth. But also Threvarthen acknowledges the transition emphasized by Tomasello, referring to it as the attainment of “secondary intersubjectivity”. As was the case with imitation, joint attention has its predecessors (Butterworth and Jarrett 1991), but it is only toward the end of the first year that children begin to participate in a “referential triangle” with an adult and an external object or event, where both are aware that they are sharing attentional focus. A theoretical generalization of these findings can be stated if we assume, largely following Tomasello (1999), that the underlying development in all these cases is the understanding of intentionality, motivating the child to do what the adult does for the reason of achieving the same goal as the adult: paying attention to “the same event”, reacting in “the same way” to a new person, performing “the same action”.

However, it need not be as suggested by Tomasello that this is done (solely) in the Inside-Out manner of “mental simulation”: the child first gains an understanding of its own intentions, and then projects these to other beings who look and behave similarly. The causality can run in both directions, as suggested by Baldwin (1913): “My sense of myself grows by imitation of you and my sense of yourself grows in terms of myself”, or as I have tried to elaborate (Zlatev 2000), in the form of a “loop” involving (1) increased objectification of one’s body, (2) increased volitional control, (3) increased “theory of mind” for others and oneself, grounded in the “bridge” of primary intersubjectivity provided to a large extent by the innate imitation capacity (Meltzoff and Gopnik 1993).

Though it still remains in question how it is achieved, it is quite clear that one of the crucial features of mimesis – awareness of one’s body – has already emerged by the end of the first year. What about representation? In Piaget’s (1952, 1953) “classic” theory of the development of representational ability (which unlike many modern uses of the term “representation” implies accessibility to consciousness), this is achieved at a later age
towards the end of the “sensorimotor period”, approximately around 18 months. By present data, this seems to be an underestimation of the child, conditioned both by Piaget’s limited evidence and by his general model.

But an important similarity with more modern developmental theories (such as those mentioned earlier) is that Piaget too considered imitation as pivotal for the origin of the first “symbols”, iconic mental representations schematizing instances of sensorimotor activity. It would be much more consistent with current terminology to refer to these as mimetic schemas. The developmental scenario proposed is furthermore surprisingly reminiscent of Vygotskyan internalization: first the child externally (overtly) imitates the actions of another person, followed by imitation of oneself in skill development, followed by the internal (covert) imitation of an inner model in imagistic thinking. Both external and internal imitation can be not only of concrete actions, but of other events perceived as similar, thus leading to mime. So despite his strongly individualistic approach to the origin of representation, Piaget also attributed a central role to mimesis in the transition to “signs” and language, and his developmental trajectory: imitation > skill > mime > mimetic schemas is consistent with the modern evidence, as well as the evolutionary progression suggested in Section 2.

3.3. Mimetic gestures and “baby signs”

Even if the child would be capable of developing mimetic schemas for representational thought without engaging in interpersonal communication, as implied by Piaget – which is doubtful (but empirically untestable) – it is clear that being embedded in a culture where such mimetic schemas are used referentially, i.e. in interpersonal communication, would boost the child’s mimetic skills even more. As mentioned earlier, infants of 9+ months begin using mimesis for intentional communication, where the most salient communicative gesture is pointing, and a few others such as waving “bye-bye”, “come” or “no” which are acquired spontaneously. But Acredolo and Goodwin (1996, 2000) have shown that from the age of 9 months (again!) children are capable of learning many so-called “baby signs”, motivated gestures, associated with a conventional meaning by parent and child, enabling the two to communicate with shared concepts even before the children can acquire language:

> 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 Goodwin 2000: 84)

The exact shape of the signs of course does not matter, but it was shown that motivatedness helps children acquire them, i.e. the system is one of (mostly) iconic gestures. Acredolo and Goodwin also present evidence that children taught “baby signs” are faster to acquire vocal language and even have a higher IQ than peers at the age of 8. This has raised both controversies and something of an industry for teaching baby-signs in middle-class families, but from an epigenetic perspective it is little surprising that
culturally stimulating a “natural” capacity for developing representation and reference would have beneficial effects on development. Furthermore, it constitutes evidence that mimetic and subsequent linguistic representations are related, as held by Piaget (1952), rather than constituting separate “layers” in the “hybrid mind”, as sometimes suggested by Donald (2001).

3.4. “Symbolic play”, differentiation and conventionalization

As noted, Piaget (1952, 1953) considered the “symbolic function”, or representational ability, developing during the second year of life to be essentially an individual achievement, something that has been strongly challenged by pointing out the strong early influence of social mediation (Vygotsky 1978, Nelson 1996, Tomasello 1999) on the infant’s mind. But even Piaget held that the “symbols”, or rather the mimetic schemas, of the late sensorimotor period were semiotically deficient in at least two ways: they were not fully differentiated from their referents, and they were idiosyncratic. Therefore they were not completely adequate both for either thinking and for communication.

It is characteristic that at this stage of development, referred to by Piaget as the “pre-operational period” children become increasingly involved in symbolic play, or pretence, both with objects and with other children. Furthermore, these interactions display a rising degree of complexity and abstractness:

Between 18 months and 2 years of age, make-believe becomes more detached from the real-life conditions associated with it. In early pretence toddlers use only realistic objects – for example, a toy telephone to talk into or a cup to drink from. Around age 2, use of less realistic toys, such as a block for a toy receiver, becomes more frequent. Sometime during the third year, children can imagine objects and events without any direct support…. (Berk and Winsler 1995: 55)

Thus, symbolic play can be seen as a “mini-laboratory” in which children experiment with representations, and in the process both learn to distinguish these from reality, how to make them more abstract, and (if necessary) to internalize them. Since these are representations of activities, we can safely conclude that they are wholly, or largely, mimetic. What is the relationship between these and language? Many studies have shown a correlation between stages of the development of symbolic play and early linguistic development (cf. Nelson 1996). The causal relationship in rather unclear, but it is natural to suppose bi-directional interaction, since in both cases (symbolic play and language) similar processes of differentiation and conventionalization of the child’s representations are played out.

Another similarity between the structure of symbolic play and language, pointed out by Vygotsky (1978), is that both require the mastery of social, normative rules which need to be internalized, i.e. the child not only acts in certain way, but understand that it should
act in that way. While it has been questioned whether children’s early language displays such awareness of normativity, it is clear that when children are capable of playing a game with roles and symbolic objects (either alone or together with others) they must be, in some sense, following normative rules: “Whenever there is an imaginary situation, there are rules.” (ibid: 95). Therefore, it is probable that mimesis is (again) leading the way, and language is following. This can also be observed in older children, when the level of linguistic competence is not sufficient for a specific function. For example, Evans & Rubin (1979) show that kindergarten children use gestures instead of speech in explaining game rules, whereas the gestures of five- to ten-year-olds are much more redundant in relation to their verbal expressions (reported by Gullberg 1998:75).

3.5. Pedagogy: ZPD and “scaffolding”

Not only language, but also pedagogy has been convincingly argued to be a distinctly human trait (Premack 1987, Donald 1991, Tomasello 1999) and it is not difficult to see a connection even though language is not literally “taught” to children. An influential concept in explaining the way in which pedagogy empowers an individual’s cognitive development is that of the Zone of Proximal Development (ZPD) defined by Vygotsky as:

…the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers. (1978: 26)

As similar idea is expressed by the concept of “scaffolding” provided by the adult in support to the developing child, and gradually “removed” as the child’s competence increases. A problem with this metaphor is, though, that it is too static to do justice to the highly dynamic and adaptive interaction between parent and child, described by Berk and Winsler (1995) as follows:

A useful analogy for the complex interaction that occurs during scaffolding is an intricate dance between a teacher (the adult) and a pupil (the child), with the child leading and the adult following for instructional purposes. (ibid: 29)

The alternative figure of a dance highlights two different ways is which the ZPD/scaffolding phenomenon may be said to involve mimesis. The first and more literal is when a manual or bodily skill is being transmitted, requiring imitation learning. The second is more general and concerns:

…the ability of the adult to sense what the child can, and can not learn – in other words to judge the ZPD. This skill is also mimetic, in as much as it involves developing a “model” of the child’s current knowledge. (Donald 1991: 177)
While one might accuse Donald of overextending the concept of mimesis here – since it is not clear how this involves the employment of representational bodily skills – the suggestion is worth looking into. The emphasis on “sensing”, rather than “theorizing about” the abilities of the child is more phenomenologically true to the experience of day-to-day pedagogy, and could help explain its highly dynamic, context-sensitive character. A hypothesis along these lines would be that the neural structures involved in imitation, “mentalizing” (Firth 1989) and “ZPD-sensing” would significantly overlap.

3.6. Summary and conclusions

In this section, I reviewed evidence for a number of related phenomena in child development, showing both a partially innate predisposition and an important mediating role for mimesis:

- infant imitation
- the rise of intersubjectivity
- the development of mimetic schemas
- iconic “baby signs”
- symbolic play
- ZPD and scaffolding

Their role is “mediating” in the sense that they constitute epigenetic prerequisites, “stepping stones” in development, without which further development would be if not blocked, at least seriously compromised. Though some of the relationships are still conjectural, it seems likely that without imitation (capacity), neither intersubjectivity nor the representational capacity evident in mimetic schemas would develop. On their side, mimetic schemas along with symbolic play are instrumental to making individual representations more conventional and normative, and to the meta-representational awareness necessary for language. The teaching of “baby-signs” is of course not necessary, but it shows an ability to understand iconic representations from as early as 9 months (and seems to have a positive effect on the development of language). Finally, pedagogy involves mimesis through imitation and possibly through the on-line “modeling” of the pupil’s abilities.

But is there conclusive evidence for a mimetic “stage” in ontogeny? The discrepancy between the time periods assigned by Nelson (1996) and Zlatev (2001a) for such a stage show that the whole concept may be problematic. Both authors define the type of stage though “a particular type of representation … dominate[ing] during a particular period of ontogenetic development” (Nelson 1996: 86). But there is a great deal of arbitrariness in deciding which representation “dominates”. The evidence reviewed in this section shows that while Nelson has underestimated early mimesis, at least from 9 months onwards, in my previous work, I had simplistically assumed that when language “gets off the ground”

3 Corresponding to the concept of “symbols” (Piaget 1962), and possibly that of “image-schemas” (Johnson 1987, Lakoff 1987).
around the age of 2, mimesis will automatically be assigned a secondary role. The simple compromise would be to merge the two proposals and extend the mimetic stage from 9 months to 4 years. But Nelson offers no hard evidence that language is used only “pragmatically”, rather than representationally prior to the age of 4. The early language-dependency of gesture (McNeill 2000) rather suggests the opposite.

The “stage” concept in ontogenesis has in recent years come under increasing critique for similar reasons. An alternative metaphor proposed by one such critical account is that of Siegler (1996) – children’s thinking develops in the form of partially overlapping “waves”. This concept may do more justice to the evidence presented in this section, as well as to the concept of mimesis itself. The point is to re-conceptualize the epigenetic role of mimesis not as a “stage” but rather as waves of mimesis, preceding waves of symbolization in a life-long perspective. Instead of mimesis “giving way” to symbolization, it rather starts before it, and makes it possible by providing various supporting structures. When symbolization, in particular language, rises – the two start a dance, with mimesis serving as a backdrop, and taking precedence whenever necessary, such as when a new activity is encountered (kindergarten games or a new language) or when language is disrupted (aphasia). Such a view saves us from the need to determine an arbitrary point in time when mimesis is no longer “dominant”.

4. Mimesis: a necessary precursor to language?

After having discussed whether mimesis may have served as the “missing link” between ape and human cognition and communication in evolution, and whether it helps the pre-linguistic child enter the symbolic world surrounding her in the previous two sections, I wish to consider briefly whether language may be not only contingently dependent on mimesis, but necessarily so. By this I mean a kind of combination of conceptual and natural necessity as will be exemplified below. At least three different aspects of mimesis can be (and have been) presented as candidates for playing such a role: its role for skill development, for iconic representation and for intersubjectivity.

4.1. Mimesis as skill

Considering the broad definition of mimesis involving the “volitional control of the body” and the fairly uncontroversial claims that (a) no complex skill can be learned without such control and (b) language consists at least in part of such complex skills, it follows logically that mimesis is a prerequisite for the learning of language, either as speech, Sign or writing. This rather obvious conclusion is often overlooked in theories of (child) language acquisition, and is therefore correctly pointed out by Donald:

Mimesis is logically prior to language, because without it, we cannot rehearse or refine any skill, let alone one so complex as speech or language. (Donald 2001: 268)
But the general way in which mimesis-as-skill serves as a prerequisite for learning any complex motoric skill makes it rather uninteresting as far as the specificity of language is concerned. There is after all, nothing semiotic per se in motoric skills, while the crucial property of mimesis is its representational character. Furthermore, it was suggested that both in phylogeny (cf. 2.5) and ontogeny (3.2) the self-consciousness necessary for volitional body control may itself be dependent on another more fundamental mimetic aspect: (proto-)imitation.

4.2. Mimesis as iconic representation

A second route to seeking a necessary connection between mimesis and language is the representational character of language: the fact that one important function of language is to serve as a “picture of reality” including both iconic (mostly structural) and arbitrary (mostly phonological) features (cf. Itkonen, this volume). But the question is if iconicity is a necessary feature of language, and if mimesis is a necessary road to understanding the concept of representation. The answer to both questions must be “No”.

The representational function itself could indeed be argued to be necessary – even though this is denied by (strong interpretations of) “pragmatic” theories such as those of the late Wittgenstein (1953) and Austin (1962). Unless a language contains the possibility to be about some state of affairs, rather than just being “associated” with it, or triggering some behavior “appropriate” for it, we would hardly call it language. But the mapping between the component parts of the utterance and the represented situation can in principle be completely arbitrary (even if it is not so in practice), and therefore iconicity can not be considered necessary.

It is also impossible to insist that iconicity of the type characteristic for mimesis is a necessary step for understanding the concept of representation, even if this turns out to be the actual situation in both evolution and ontogeny. For one thing, experiments with teaching symbol systems to apes have shown that they are capable of learning a limited number of conventional meanings of symbols, whose form is arbitrary in relation to their referents (Savage-Rumbaugh and Rumbaugh 1993, Deacon 1997).

In sum, even if mimesis possesses advantages compared to arbitrary symbol systems in terms of transparency and ease of memorization, which motivates its early appearance in both evolution and ontogeny, these are not strong enough to claim that it is unthinkable that language could emerge without it.

4.3. Mimesis as a ground for intersubjectivity

A final, and I believe strongest of the three arguments for mimesis playing a necessary role for the emergence of language lies in the socially shared nature of linguistic meaning (Wittgenstein 1953, Itkonen 1983, Nelson 1996). This implies not just that the meaning/concept of expression X “inside” Speaker A and Speaker B happens to be the same, or that A’s and B’s individual concepts becomes somehow “calibrated” with usage (Steels, Kaplan, McIntyre and Van Looveren 2002), but that A and B have an active way
of “negotiating” the meaning of X so that it is the same for both of them, and they know that it is the same. The way in which this is done is not too different from the way A and B can determine that they are looking at the same object, i.e. participate in acts of joint attention.

This last comparison suggests where the necessary connection between mimesis and language may lie: joint attention rests on mimesis in so far as it involves the imitation of purposeful action. Imitation (in a general sense), whether it rests on a partially innate “mirror neuron system” or not, is very likely the only way for developing structures of shared experience such as joint attention. In that case, without mimesis there would be no intersubjectivity, and without intersubjectivity, there could be no language. It is hardly conceivable how language itself would enforce such structures of shared meaning – since without non-linguistic, at least in part mimetic criteria of successful communication, language would be unlearnable. In the absence of miraculous factors, mimesis stands as a necessary prerequisite for language. This, I believe, is the meaning behind Donald’s somewhat cryptic statement:

Language is different from mimesis, but is 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)

In emphasizing that language always occurs in consensual “games” or “forms of life”, Wittgenstein (1953) seems to have made a very similar point. What I am here suggesting is that mimesis is necessary for the establishment of these games, i.e. that they are originally “mimetic games”, prior to becoming “language games”. The failure to engage in such games typical for autistic children (Firth 1989) unsurprisingly has detrimental effects on language.

5. Conclusions

The discussion in this paper has shown the potential of the concept of mimesis, defied as by Donald (1991, 2001) as conscious, representational body motion, for explaining the “roots” of human language are considerable. In being consistent with a large amount of evidence from human evolution and ontogeny, it serves a unifying role. Conceptually, it can help us understand the formation of shared experience, and ultimately shared symbolic meaning, which is perhaps the central characteristic – and puzzle – of language: if concepts are psychological entities, how can language be social?

At the same time, one conclusion is to refrain from uncritically applying the concept of a “mimetic stage” as a bridge between the signal-based communication of animals and pre-linguistic children – on the one hand – and the symbolic language of adult human beings on the other. For one thing, mimesis was shown not to be a unitary concept, but to

---

4 Save for e.g. telepathy.
involve (at least) the phenomena of mime, imitation, skill and gesture some of which are inherently social, others are not, some are explicitly representational, others are not. An urgent task is to attempt to reconstruct the order in which they have emerged in evolution, which may also parallel their development in childhood. Largely following the Vygotskian principle of interpersonal-to-intrapersonal development, a likely progression is the following:

proto-imitation > imitation > (panto)mime > mimetic schema > gesture

What characterizes this order is that it coincides with increased semioticity, i.e. differentiation of expression from content and an explicit representational relationship, abstractness and internalization. Proto-imitation, both in phylogeny and ontogeny does not involve an understanding of the model’s goals, while imitation does. Mime is more rather than less complex, since its function is purely representational rather than pragmatic. Mimetic schemas are essentially internalized and schematized mimes, and correspond structurally to what cognitive linguists call “image schemas” (Johnson 1987) – but a central difference is that since they are based in social rather than individual experience, they stand in a better position to “ground” linguistic meaning than purely sensorimotor structures. Gestures involve sequences of mimes used for communicative purposes, differing in degree of conventionality and language-dependence (Kendon 1988).

In sum, there is considerable empirical support for the hypothesis that mimesis has preceded and made possible symbolic language in evolution, and continues to do so in ontogeny. But at least in the latter case, this is more in the manner of “waves” serving as support structures for the development and use of language, rather than a “stages” stable enough to be self-contained. Much of the argumentation presented in this paper has been conjectural, and there are many questions remaining, but with its potential for filling in gaps in our understanding of semiotic development, for integrating evidence from phylogeny and ontogeny, and bringing in both empirical and conceptual issues, the mimetic hypothesis is certainly worth pursuing further.

References


