# Life after the symbol system metaphor*

Karl F. MacDorman
School of Informatics, Indiana University

After reviewing the papers in this special issue, I must conclude that brains are not syntactic engines, but control systems that orient to biological, interindividual, and cultural norms. By themselves, syntactic constraints both underdetermine *and* overdetermine cognitive operations. So, rather than serving as the basis for general cognition, they are just another kind of empirically acquired constraint. In humans, symbols emerge from a particular sensorimotor activity through a process of *contextual broadening* that depends on the coordination of conscious and nonconscious processing. This process provides the representational freedom and stability that constitute the human brain's solution to the frame problem and symbol grounding problem. Symbol formation and grounding is an ongoing process of generalising constraints from particular contexts, selectively enlisting their use, and re-automating them. This process is central to the self-creation of a language-using person with beliefs, agency, and identity.

**Keywords:** affordance, conduit metaphor, distributed cognition, embodiment, language of thought, the symbol grounding problem

## Introduction

Those writing in this special issue concur that "human interaction and thus human symbols cannot… be reduced to the formal units pinpointed by syntactic analysis" (Belpaeme and Cowley, 2007). In spite of this, nobody draws the logical conclusion. We should stop conceptualising symbol grounding in terms of formal symbol systems. Modelling can proceed without treating 'minds' as syntactic engines. Therefore, I believe we need to rethink the *person problem*:[1] How can human bodies — and perhaps robot bodies — construct themselves into persons by attuning to patterns and norms in their social environments? I believe an approach that addresses this question could satisfy the authors' conviction that, to engage us in talk, robots will need to do more than ground their internal symbols in the world.

*Interaction Studies* 8:1 (2007), **143–158**.
ISSN 1572–0373 / E-ISSN 1572–0381 

For 'walking, talking' persons, operations are dynamic (Thelen & Smith, 1994), and what we do is irreducible to the formal and syntactic (Clocksin, 1998). The ability to use (external) language is rooted in the body, with brains acting as control systems. So, how do human beings accomplish this? Broadly speaking, there seems to be some kind of consensus among the authors. Using the physics of speech (Worgan & Damper, 2007), human beings align to regularities that include patterns of use (Vogt & Divina, 2007). This enables the later development of what can be *described* as symbolic reasoning (Clowes, 2007). Initially, at least, we are not 'internalising' symbols but using them as constraints on what we do (Clowes, 2007; Cowley, 2007; Viger, 2007). What we characterise as symbol use (or word use) is the outcome of a culturally-located developmental process.

Researchers in artificial intelligence and robotics have often been reluctant to recognise the magnitude of the task they face. One reason for this is that it is easy to be led astray by the *conduit metaphor* (Reddy, 1979). This metaphor characterises, not the nature of language, but how we talk about it. According to this view, a speaker puts ideas into word-containers that are shunted along a conduit and then taken out by a hearer. Cognitive scientists have elevated this 'putting' and 'taking' with such terms as 'encoding' and 'decoding,' but the basic idea of translating concepts from one head into the local language and then back into someone else's head remains the same.

Given the allure of the conduit metaphor, many have refused to recognise the difference between the problem of grounding *internal* symbols in perceptuo-motor invariances as confronted by formal symbol systems (Harnad, 1990) and the *external symbol grounding problem,* namely, the problem of grounding the *external* symbols and signals of utterance activity. This distinction is crucial, because the evolving cultural resources that shape *human* cognition include these external symbols and signals (Lyon, Nehaniv & Cangelosi, 2007; Vogt & Divina, 2007). External symbol grounding depends on real-time co-action (Cowley, 2007), which includes teaching (Seabra Lopez & Chauhan, 2007), and prompts the rise of perceptual skills (Steels & Belpaeme, 2005). Thus, human beings face not only the challenge common to other species of keeping internal representation grounded in the world but also a quite different challenge of grounding (external) language. Damper and Worgan, Cowley, Clowes, Viger and others are correct in maintaining that, although we begin with the physical properties of utterances, it is only through a developmental process that human beings are gradually induced to discover the power of their formal properties.

Although we talk about language as if it were a conduit, this metaphor blinds us to how agents can use representations in turning themselves into people. This developmental process occurs as representations gradually acquire symbolic properties. When symbolic ways of acting are consistent with an agent's developmental history, they give rise to representations that can stand for objects, including objects that are out of view (Viger, 2007). The ability for symbols to act as stand-ins beyond the context of their initial use is typically viewed as *decontextualisation.* This, however, is to substitute a logical for a developmental point of view. From the perspective of grounding, what happens is more appropriately seen as *contextual broadening*. Thus, to solve the *person problem*, we can simulate how simple causal regularities give rise to contingencies that enable systems to reconfigure embodied representations, including those that depend on the movement of other bodies (Cowley, 2007). This involves simulating the interplay of conscious and nonconscious processing,[2] which has been scandalously ignored by past approaches to symbol grounding (MacDorman, 1999). So, in robotics and artificial intelligence, we have to move from forms of representation that can only be reconfigured by human programmers to forms that have sufficient representational freedom and stability to reconfigure themselves based on the interplay of conscious and nonconscious processing. Nothing else will give an agent symbols it can use with sensitivity to norms to constrain the doings of both its own body and those of other agents (Cowley & MacDorman, 2006). This will provide it with the robust grounding that is required to solve the person problem.

## Persons are not formal symbol systems

Just as vacuum tubes and other technologies eventually reach an end of life, so do ideas. Symbol systems as defined in Harnad (1990a) seem to have reached just such an end, because, by definition, they can only operate according to *syntactic* constraints. If other constraints (e.g., perceptual, ecological, biomechanical, epigenetic, evolutionary) are to influence a symbol system, it will be solely by virtue of how they connect its symbols to the objects, events and relations they represent.[3]

However, *if* persons (or their brains) were symbol systems, their cognition would operate according to constraints that were simultaneously biological, interindividual, and cultural. In short, they would *not* be purely syntactic. *Our brains are not syntactic engines but world-oriented control systems* and, as Vygotsky (1986) saw, follow different *lines* of development in dealing with physical and cultural/intentional entities. It is quite possible that verbal thought and meaning become systematic, generative, and inferentially coherent as a consequence of self-organising brains aligning us as persons to cultural norms (MacDorman & Cowley, 2006). These regularities serve as standards against which we evaluate each other's behaviour from our own perspective and, by doing so, give that behaviour value and meaning (Christensen & Bickhard, 2002; Cowley & MacDorman, 2006). These capacities develop in the affectively rich relationships of infants and their caregivers (Cowley, 2007). Thus, the rule-like regularities of human cognition do not reveal syntax as the underlying, built-in mechanism that governs symbolic reasoning.[4] Rather, these regularities emerge from causal processes that depend on a language-saturated cultural ecology. We exploit both the physical world and sets of categories that are derived from the history of particular cultural domains. As Dennett (1987, 1989) and Ross (2006) argue, we gradually come to use the narrative devices that turn our human bodies into persons.

Fodor (1980) and others have proposed that a syntactically-driven symbol system has *a priori* feature detectors for every natural kind of thing. These detectors set processing in motion and, thereby, instantiate the symbols of the symbol system. An example of this from artificial intelligence is the robot Shakey (Nilsson, 1984). From the standpoint of evolution or neurobiology, however, *a priori* feature detectors do not seem plausible (MacDorman, 1998), as even Fodor admits (Guttenplan, 1994). Those who would fix Good Old Fashioned Artificial Intelligence (GOFAI) misunderstand Harnad (1990a), if they believe the matter of symbol grounding only concerns whether a robot is hardwired with feature detectors or learns them on the fly. Consequently, some have simply proposed to connect a pre-existing symbol system to the world by means of a perceptual learning mechanism, such as a neural network.[5]

This setup spares the old *syntactically-governed* symbol system idea, while appearing to address its grounding. But it misses the force of Harnad's point — that, just as a diagram may guide someone in performing formal geometry, the iconic shape of a symbol's referent must constrain the system in *manipulating* the symbol. Thus, Harnad argues that, *to ground symbols, empirical constraints must augment syntactic constraints.* It is not enough that the pattern recognition mechanism by which symbols are instantiated be learned. Constraints on the rules or rule-like regularities by which symbols are manipulated must also be learned. The reason for this, I would argue, is because ecological relations and experience (and, in human beings, norms and language) influence not only how we recognise what is around us, but also how we reason about it (Wason, 1981; MacDorman, 1999). Fodor (1987, 2000) opts for a similar approach in his solution to the philosopher's version of the frame problem: Empirical constraints need to augment syntactic constraints to eliminate "kooky concepts" and thereby help the system avoid reasoning about things that cannot happen or things that do not change.[6]

Both Fodor and Harnad assume that the main problem with syntactic constraints is that they underdetermine mental representations. In other words, syntactic constraints define too broad a universe of possible mental representations, so that an individual's actual mind is an empirically-delimited subset of that universe. This makes Harnad the 'Gorbachev' of cognitivism: He tries to save the approach while fatally undermining it. He considers formal symbol systems to be something that can be fixed by supplementing syntactic constraints with empirical ones. It is as if, like Fodor, he thinks symbol systems are the only way of making behaviour systematic, generative, and inferentially coherent. This view, however, is mistaken. The *a priori* syntactic constraints of symbol systems not only underdetermine mental representations *but, just as importantly, overdetermine them* by excluding nonsyntactic mental operations. In human beings, syntactic constraints may themselves be a product of cultural evolution (Vogt & Divina, 2007). Moreover, rather than apply the formal symbol system model to the whole of general cognition, one may posit that other kinds of systems supply the richness and parsimony required to represent what the brain needs to manage action. We cannot simply assume symbol systems are the building blocks with which body and brain construct a person. On the contrary, these patterns are likely to exist in the world. Thinking may appear systematic, generative and inferentially coherent, because we use social practices to abstract these patterns from the experience of living in a culture. They get the job done and are socially sanctioned. The historically recent development of writing systems, formal education, and computers have further reinforced the dubious belief that meaning derives mainly from the compositionality of symbols (Luria, 1976).

So, if mental operations are neither formal nor syntactic, the issue of symbol grounding should concern how human bodies construct themselves into walking, talking persons from their social environments. It should concern not just the learning of perceptual invariants but also the simulation of persons who can follow (or break) norms and rules. Thus, we are interested in how agents can be educated to think in ever more (or less) logical ways and how they can develop stories to 'explain' why they do what they do, taking an intentional stance toward themselves and their actions (Dennett, 1987, 1989; Ross, 2006). Through this narrative process, people spin explanatory 'myths.' Indeed, in recent decades people have convinced themselves that cognition is driven by processes that are like a telephone exchange, or a computer, or whatever happened to be the dominant metaphor of the day.

## The conduit metaphor obscures language and cognition

The conduit metaphor (Reddy, 1979) tempts us to compare perception with communication. In both cases, we assume, the world provides information that we decode: For perception, this decoding is dependent on sensory transduction, perceptual inferencing, and modelling (Ullman, 1980). For communication, it depends on using an inner system (a language faculty) to decode a signal in terms of invariants that have determinate value (Jackendoff, 2002; see Port & Leary, 2005, for an opposing view). But the metaphor of carrying information distorts our view of how language is learned, how it functions interpersonally, and how it shapes social cognition. There has been an opinion in some branches of mainstream cognitive science that essentially goes like this:

1. Human beings think in a formal language of thought (LOT), which constitutes the symbolic representations of a symbol system (Fodor, 1975).
2. Operations in LOT depend only on internal syntactic constraints.
3. Nevertheless, LOT maintains semantic correspondences with the external world: When you feed true statements into a LOT-manipulating symbol system, the system produces conclusions that by and large are also true.
4. Since the symbol system operates according to internal syntactic constraints, elementary symbols are innate, given beforehand.[7]
5. Complex representations are composed from an 'alphabet' of these symbols.

According to this viewpoint, communication entails encoding in natural language (e.g., an utterance or email) mental concepts, which are taken to be symbolic representations, and sending them across a conduit (e.g., airwaves or the Internet) to a recipient who then decodes them.[8]

The conduit metaphor comes up short for a number of reasons:

1. For us to communicate, your concepts do not need to divide up the world in exactly the same way as my concepts (Reddy, 1979; Steels & Belpaeme, 2005).
2. There is much more to communication than the transfer of facts about the world expressed as propositions.
   a. Kravchenko (2003, 2006), following Maturana (1976) and Maturana and Varela (1980), argues that human communication — like that of any other animal — must be connotational.
   b. As Cowley and MacDorman (1995) have shown, the descriptive content of an utterance often says little about what is happening between individuals in a relationship and what the utterance means for them. An understanding of prosodic, facial and gestural features of communication as well as context is needed.
   c. As Karl Grammer, Alex Pentland, and others point out (e.g., Grammer, Fieder & Filova, 1997; Pentland, 2005), communication can be conceived of in many ways other than in terms of semantics (descriptive meaning) and pragmatics. For example, during human courtship, nonverbal signals have a probabilistic quality that can give information about receptiveness to a romantic advance without making a firm commitment, thus allowing 'wiggle room' to back out.

3. There is much scepticism about the internalist assumptions of Fodor's LOT, and many researchers follow Dennett (1991), Clark (1997), Hutchins (1995) and others in viewing language as a process of enacting external patterns.[9] While the idea remains underdeveloped, we construct ourselves in the course of experiencing the external social environment (e.g., Dennett, 1987, 1989; Ross, 2006).

The physical symbol systems hypothesis and its problems are symptomatic of people's belief in the conduit metaphor (Newell, 1980).

Given that symbol systems function according to internal syntactic constraints, how do we ground their symbols in the external world? I have argued for more than 12 years that the whole idea of a system operating according to internal *a priori* syntactic constraints is wrong. Rather, the system's rule-like operation and all its symbols must be learned from the bottom up. They must emerge from and be grounded in sensory projections, motor actions, and affective consequences. The system's operation turns on empirical constraints of which formal syntactic constraints are no more than a subset.

In the early 1990s, I was approaching the symbol grounding problem from Stevan Harnad's viewpoint. Cowley (1994, 1997) was concerned about something else: how features of utterances operate between individuals to closely coordinate their activity and regulate their emotion in real-time. These features can be interpreted in many different ways and, based on his acoustic data from conversations, Cowley saw no reason to privilege text-based 'symbolic' interpretations over those that arise from our acute sensitivity to closely-timed prosodic events (e.g., pitch, rhythm, loudness, voice quality). Cowley was (and remains) very concerned about how language gets bootstrapped by the mother-infant dyad. For example, when a baby reaches for a stuffed dog and makes a vocalisation, the mother often over-interprets this as asking for the dog by pointing at it, and the baby cognitively grows into this over-interpretation through the mother's coaxing (see Tomasello, 2003).

So the question Cowley was asking is: How is it that an utterance that starts off with no meaning comes to serve various functions within the mother-infant relationship and eventually acquires semantic meaning? It may, for example, come to act as a stand-in for an object that is not present (Viger, 2007). Because the features of the utterances, including its symbolic/representational features, are in the environment, I suggested to Stephen Cowley, "That's not the *internal* symbol grounding problem, but the *external* symbol grounding problem." By that I meant, "That's not the problem of how you ground the internal symbols in your head in the external objects, relations and states-of-affairs they represent (as formulated by Harnad, 1990a). Rather, that is the problem of how external utterances and features of utterances (including, but not limited to, symbolic features) can come to regulate affect and coordinate activity between persons (e.g., mother and infant) and eventually stand in for things not present." It is this idea that underlies Cowley's attempts to understand how human infants deal with symbol grounding (Cowley, 2006, 2007; Cowley et al. 2004).

So we have three problems: (1) the problem of grounding internal symbols in external states-of-affairs; (2) the problem of grounding utterances (including prosody and the actual words spoken) and facial, gestural and eye movements in interindividual activity (including affect); and (3) the problem of using robots to investigate not only the first symbol grounding problem but also the extended version having to do with human language.[10]

## How brains ground symbols

Being part of a symbol system, Viger (2007) argues, is not the defining characteristic of a symbol but rather its potential to stand in for something that is out of view. But a symbol's potential to function in *different contexts* is equally important, because this is how an agent shows it understands a symbol *as a symbol* and not just as a pattern. This potential is often called *decontextualisation*. While Viger (2007) acknowledges its importance, he notes that it presupposes that the symbol comes first. However, because symbols presuppose an interface, it is more appropriate to posit that the grounding comes first and that the symbol instead emerges from a process of *contextual broadening*. When a baby reaches for a plate and says, "mo," and the mother interprets this to mean, "I would like to have some *more* food," this does not mean "mo" functions as a symbol for the baby. The baby may just be repeating the mother's vocalisation (or its own). It may learn that in this context, vocalising "mo" is a way of acting that gets the mother to approach with food. Only when the baby exhibits an understanding of how "mo" (or "more") functions in different contexts can we say that it knows the vocalisation stands for a particular abstract relation (e.g., between how much it has and how much it wants). In this sense, the baby demonstrates a grasp of the word's meaning through use (Wittgenstein, 1953).

Conscious processing may be implicated in contextual broadening, because persons can think about objects in focal consciousness in ways that indicate the loosening of the objects from context. For example, I can think about binding papers with a stapler or using it to put up posters, but I can also think about the effects of microwaving it, or things I cannot do with it, like throwing it into the sun. This indicates a degree of *representational freedom* for objects in consciousness that might *not* be obtained by strictly nonconscious processing.

But nonconscious processing has its advantages too. It supports the parallel execution of well-honed skills. The reason a FIFA professional can play soccer is because he *doesn't need to think* about how to run or dribble or make a shot. While these things are coming together automatically, he can focus on strategy. If you were to make him think about, for example, ways to improve his shooting, presumably his game would worsen before it improved (Langer & Imber, 1979). Like soccer, learning to drive takes conscious study and effort, but eventually people not only manage to drive but to do many other things while driving. Yet, if something happened on the road, extraneous activities would be interrupted as orienting reactions redirected conscious attention to the unexpected event.

It is precisely because conscious processing is so useful in evaluating novel stimuli that it must be supplemented by other mechanisms. It typically has widespread effects on attention, memory, and motor control (Baars, 1988). While contradictory beliefs may coexist in the subconscious, objects in consciousness are required to maintain a much higher degree of consistency. For example, a locked gate forced me to park my bicycle in a new place, but later without thinking I returned to its usual spot only to find it missing. Outside of consciousness, my brain was quite content to represent the bicycle as residing at its usual location and at its new location. Flexibly maintaining consistent representations in consciousness may be computationally demanding, involving billions of bits of information in the brain being simultaneously and coherently interrelated (MacDorman, 1999, 2004). This may explain why the brain's massively parallel nonconscious processing can break down into serial processing in consciousness (Mangan, 2001). This is not to deny the capacity of nonconscious processing to maintain complex interrelations that have already been worked out through practice. But it is this "working out" that is the domain of conscious processing, because what is represented consciously is articulated with sufficient richness to allow persons to evaluate novel situations.

If solving the *frame problem* entails finding a representational form that allows computational agents to avoid getting bogged down reasoning about stable aspects of the world (Janlert, 1996), the interplay between conscious and nonconscious processing has endowed the human organism with a solution. The brain tolerates inconsistencies until something goes wrong — something unexpected happens. It then sets conscious processing on to the problem — with its flexibility, consistency checks, and integration of disparate brain centres — to correct and reautomate the subconscious routines that led to the error. The result of this process is that human brains can deploy (1) a flexible and consistent representation of objects in consciousness, (2) a vast number of stable subconscious routines that are able to run in parallel, and (3) methods for consciously detecting when things go wrong and correcting and reautomating those routines. Isn't this what an intelligent robot needs to be able to do?

Let us now consider the two major robotics approaches to intelligent behaviour. There is the GOFAI approach of programming a robot to make plans by means of a symbol system and the behaviour-based approach of building up complex behaviours from simple mechanisms that directly link sensing with motor response. The GOFAI approach, which suffers from the usual symbol grounding and frame problems, metaphorically resembles conscious processing, while the behaviour-based approach with its hardwired routines metaphorically resembles some kinds of nonconscious processing. Some researchers have taken a hybrid approach by welding together GOFAI and behaviour-based subsystems along a fixed interface (Malcolm, 1995). But it is clear that none of these approaches attempts to simulate the interplay of conscious and subconscious processing, which is how the human organism has finessed the symbol grounding and frame problems.

Although it is easy to argue that GOFAI-based robots have too much representational freedom, while behaviour-based robots have too little, this argument really misses the point. To illustrate why, it is useful to consider Brooks' robot Herbert (1991a, 1991b). Herbert appeared to perform an intentional activity, collecting soda cans in an office, but it did so without benefit from any form of central symbolic representation. When a sensor detected a can, a switch sent the robot forward. When its collision with the desk stilled the wheels, a gripper extended. When the soda can broke an infrared beam in the gripper, it grasped the can, and so on.

Is Herbert *too grounded* in the situation and therefore lacking representational freedom? I would argue against it, and just because a person can think abstractly about soda cans does not make that person's thinking less grounded than Herbert's responses. Herbert has a brittle kind of grounding, which, oddly, is similar to the GOFAI-based robot Shakey (Nilsson, 1984). Despite their differences, Herbert, like Shakey, is crafted by engineers to perform specific actions on specific items in specific environments (e.g., where desks have a certain height). Neither robot could get very far in a new environment with unknown objects. While well-fitted by a (conscious) designer to a given purpose, neither could reconfigure its internal workings for some new task. Neither robot can discover what objects afford, nor develop and automate skills for handling objects based on their affordances (Gibson, 1979). In other words, they lack the kinds of cognitive systems that, in humans, coordinate conscious and nonconscious processing. Just as traditional symbol systems cannot be grounded without a human interpreter (Harnad, 1990a, 1990b), Brooks' subsumption architecture cannot be grounded without a human designer to rebuild it as the context changes.

Neither Shakey nor Herbert has a mechanism for *re*grounding, which explains why their grounding is so brittle. It fits a certain environment but fails irrecoverably in others. A system with a flexible grounding will reground itself when its body, environment, or even its goals change (MacDorman et al., 2000). Like the babies Cowley (2007) studies, it will be able to use the world to construct its own agency. Unfortunately, however, the apparent grounding of Herbert and Shakey depends on the plasticity of their designers' brains and not their own. And so it is their designers' brains that effectively coordinate processing that is more centralised, articulated, and conscious with processing that is more automated, modularised, and subconscious (MacDorman, 1999). In this manner, information is decontextualised and recontextualised as skills are automated and reautomated. It is these cognitive and sensorimotor processes that make for a robust grounding.

## Conclusion

Except in abstract, formal systems, such as those implemented on a computer, even symbolic processing is embodied and situated — and influenced and constrained by its embodiment and other circumstances (Pattee, 2001; Lindblom & Ziemke, 2006). A given body sets up unique ecological relations between its perceptuomotor and cognitive systems and the environment. Contextual broadening develops from the sensorimotor projections of objects in a *particular* situation. The ability to reason in other, new situations does not involve substituting syntactic constraints for empirically acquired ones but rather acquiring and generalising empirical constraints from a history of interaction and bringing them to bear on the situation at hand. In more abstract planning and reasoning, empirically acquired constraints may take on logic-like properties, but ecological (i.e., body–world) constraints remain at play (Wason, 1981). Even logical constraints are learned empirically. Indeed, logical thinking is just one of many kinds of (learned) habits that afford contextual broadening.

Fodor (1996) has characterised the symbol systems of AI as being subject to the frame problem, because they have too much representational freedom, being governed only by syntactic constraints, while Harnad (1991) has attributed the same cause to the symbol grounding problem. However, I would argue that syntactic constraints are a subset of empirical constraints and not the other way around. Symbol systems do not suffer from too much representational freedom but the wrong kind! In contrast, the human brain effectively manages the trade off between freedom and stability through the interplay of conscious and nonconscious processing, and that is our solution to the frame and symbol grounding problems.

## Notes

* Much appreciation goes to all the participants in the External Symbol Grounding Workshop in Plymouth, UK on 3 and 4 July 2006, and especially to Tony Belpaeme, Stephen J. Cowley, Stevan Harnad, Alexander Kravchenko, Joanna Raczaszek, and Christopher Viger for fruitful discussions on symbol grounding that have contributed to this paper. I would also like to thank Bruce B. Mangan for a discussion on consciousness and Robert Port for a discussion on spoken language.

**1.** This label emerged from private correspondence with Stephen J. Cowley.

**2.** Nowhere in my discussions of consciousness in this paper do I attempt to address the so-called 'hard problem' of why we are even conscious at all. I am merely looking at the *functional* relation between conscious and nonconscious processing — a relation which presumably could be implemented in machines regardless of whether these machines would really be conscious. My treatment of the hard problem appears in MacDorman (2004).

**3.** They may do so, for example, by virtue of how a mushroom instantiates a *mushroom* symbol in the head and, inseparably, how mushroom-related symbol processing elicits mushroom-directed sensorimotor actions like picking, sorting, washing, frying, and eating.

**4.** This contrasts with Soar (Laird, Newell, & Rosenbloom, 1987) and other ungrounded reasoning programs implemented on computers.

**5.** The mechanism performs category induction by learning invariant properties of an object's sensory projections to distinguish the object from confusable alternatives based on internal and external state changes (Harnad, 1987; Cangelosi, Greco & Harnad, 2000).

**6.** If after the robot has lifted a cup it has to determine whether the French Prime Minister is still in office, it is wasting time reasoning about something that is highly unlikely to have changed. If it has to determine whether Venus has entered its gut, it is wasting time reasoning about something that cannot happen. And if it has to figure out that it should *not* reason about these things, it has already fallen into the frame problem. The solution is to find a representational form that is adequately constrained so that senseless reasoning will not happen.

**7.** This has been Fodor's assumption. According to Fodor, 1980, these elementary symbols are not something low-level like edges and blobs or other basic perceptual or sensorimotor features, but natural kinds of things: emperor penguins, zebras, and so on.

**8.** I may have in my head a concept corresponding to the sentence "John saw Mary with her ex-husband." But this concept, as represented in LOT, need not resemble any natural language. (It could instead resemble, e.g., Schank and Ableson's Conceptual Dependency, 1977.)

**9.** This work builds on Vygotsky's (1986) metaphor of *internalising* what begins as external mediational means. While the term *internalisation* is much disputed, it can be uses as shorthand for processes like the move from counting on the fingers to counting in the head or from talking to others to thinking in words.

**10.** Since many animals presumably can consistently perceive relations among objects, and thus enjoy systematicity in their perceptual apparatus (as Fodor and Pylyshyn, 1988, say, "punctate minds don't happen"), I can see the problem of grounding internal symbols as also being an issue for many non-primate species.

## References

Baars, B. (1988). *A cognitive theory of consciousness.* Cambridge, UK: Cambridge University Press.

Steels, L. & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences, 28*(4), 469–489.

Brooks, R. A. (1991a). Intelligence without reason. In *Proceedings of the Twelfth International Conference on Artificial Intelligence (IJCAI),* Sydney, Australia (Vol. 1), pp. 569–595. San Mateo, CA: Morgan Kaufmann.

Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence, 47,* 139–159.

Cangelosi, A., Greco, A. & Harnad, S. (2000). From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories. *Connection Science, 12*(2), 143–162.

Cheney, D. L. & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species.* Chicago: University of Chicago Press.

Christensen, W. & Bickhard, M. (2002). The process dynamics of normative function. *Monist, 85*(1), 3–28.

Clark, A. (1997). *Being there: Putting brain, body, and world together again.* Cambridge, MA: MIT Press.

Clocksin, W. F. (1995). Knowledge representation and myth. In J. Cornwell (Ed.), *Nature's imagination* (pp.190-199). Oxford, UK: Oxford University Press.

Clocksin, W. F. (1998). Artificial intelligence and human identity. In J. Cornwell (Ed.), *Consciousness and human identity* (pp. 101–121). Oxford, UK: Oxford University Press.

Clocksin, W. F. (2004). Memory and emotion in the cognitive architecture. In D. Davis (Ed.), *Visions of mind* (pp. 122-139).Hershey, PA: IDEA Group.

Clowes, R. (2007). Semiotic symbols and the missing theory of thinking. *Interaction Studies, 8*(1).

Cowley, S. J. & MacDorman, K. F. (1995). Simulating conversations: The communion game. *AI & Society, 9*(3), 116–137.

Cowley, S. J. (1994). Conversational functions of rhythmical patterning: A behavioural perspective. *Language and Communication, 14,* 353–376.

Cowley, S. J. (1997). Conversation, co-ordination and vertebrate communication. *Semiotica, 115*(1), 27–52.

Cowley, S. J. & MacDorman, K. F. (2006). What baboons, babies, and Tetris players tell us about interaction: A biosocial view of norm-based social learning. *Connection Science, 18*(4), 363–378.

Cowley, S. J., Moodley, S. & Fiori-Cowley, A. (2004). Grounding signs of culture: Primary intersubjectivity in social semiosis. *Mind, Culture and Activity, 11*(2), 109–132.

Cowley, S. J. (2006). Distributed language: Biomechanics, functions and the origins of talk. In C. Lyon, C. Nehaniv & A. Cangelosi (Eds.), *The emergence and evolution of linguistic communication* (pp. 105–129). London: Springer.

Cowley, S. J. (2007). How human infants deal with symbol grounding. *Interaction Studies, 8*(1).

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (1989). The origins of selves. *Cogito, 3,* 163–173.

Dennett, D. C. (1991). *Consciousness explained.* Boston: Little, Brown.

Fodor, J. A. (1975). *The language of thought.* New York: Cromwell.

Fodor, J. A. (1980). On the impossibility of acquiring more powerful structures. In M. Piatelli-Palmarini (Ed.), *Language and learning.* London: Routledge.

Fodor, J. A. (1987). Modules, frames, fridgeons, sleeping dogs and the music of the spheres. In Z. W. Pylyshyn (Ed.), *The robot's dilemma: The frame problem in artificial intelligence* (Chapter 8). Norwood, NJ: Ablex.

Fodor, J. A. (2000). *The mind doesn't work that way: The scope and limits of computational psychology.* London: MIT Press.

Fodor, J. A. & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28,* 3–71.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Grammer, K., Fieder, M. & Filova, V. (1997). The communication paradox and possible solutions. In A. Schmitt, K. Atzwanger, K. Grammer & K. Schäfer (Eds.), *New Aspects of Human Ethology* (pp. 91–120). New York: Plenum Press.

Guttenplan, S. (1994). Jerry A. Fodor. In Guttenplan, S. (Ed), *A companion to the philosophy of mind.* Oxford, UK: Blackwell.

Harnad, S. (1987). Category induction and representation. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition.* Cambridge: Cambridge University Press.

Harnad, S. (1990a). The symbol grounding problem. *Physica D, 42,* 335–346.

Harnad, S. (1990b). Lost in the hermeneutic hall of mirrors. *Journal of Experimental and Theoretical Artificial Intelligence, 2,* 321–327.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Jackendoff, R. (2002). Foundations of language: Brain, meaning, grammar, evolution. Oxford: Oxford University Press.

Janlert, L.-E. (1996). The frame problem: Freedom or stability? With pictures we can have both. In K. M. Ford & Z. W. Pylyshyn (Eds.), *The robot's dilemma revisited.* Norwood, NJ: Ablex.

Kravchenko, A. V. (2003). The ontology of signs as linguistic and non-linguistic entities: A cognitive perspective. *Annual Review of Cognitive Linguistics, 1*(1), 179–191.

Kravchenko, A. V. (2006). Cognitive linguistics, biology of cognition and biosemiotics: Bridging the gaps. *Language Sciences, 28*(11), 51–75.

Laird, J., Newell, A. & Rosenbloom, P. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence, 33*, 1–64.

Langer, E. & Imber, L. (1979). When practice makes imperfect: The debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37,* 2014–2025.

Lindblom, J. & Ziemke, T. (2006). The social body in motion: Cognitive development in infants and androids. *Connection Science, 18*(4), 333–346.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations.* Cambridge, MA: Harvard University Press.

Lyon, C., Nehaniv, C. L., Cangelosi, A. (2007). Emergence of communication and language. London: Springer.

MacDorman, K. F. (1998). Feature learning, multiresolution analysis, and symbol grounding: A peer commentary on Schyns, Goldstone, and Thibaut's 'The development of features in object concepts.' *Behavioral and Brain Sciences, 21*(1), 32–33.

MacDorman, K. F. (1999). Grounding symbols through sensorimotor integration. *Journal of the Robotics Society of Japan, 17*(1), 20–24.

MacDorman, K. F., Tatani, K., Miyazaki, Y. & Koeda, M. (2000). Proto-symbol emergence. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* October 30-November 5, 2000. Kagawa University, Takamatsu, Japan.

MacDorman, K. F. (2004). Extending the medium hypothesis: The Dennett-Mangan controversy and beyond. *Mind and Behavior, 25*(3), 237–257.

MacDorman, K. F. & Cowley, S. J. (2006). Long-term relationships as a benchmark for robot personhood. In *Proceedings of the 15th IEEE International Symposium on Robot and Human*

*Interactive Communication* (RO-MAN). September 6–9, 2006. University of Hertfordshire, Hatfield, UK.

Malcolm, C. M. (1995). The SOMASS system: A hybrid symbolic and behaviour-based system to plan and execute assemblies by robot. In J. Hallam et al. (Eds.), *Hybrid problems, hybrid solutions* (pp. 157–168). Oxford: ISO Press.

Mangan, B. B. (2001). Sensation's ghost: The non-sensory 'fringe' of consciousness. *Psyche, 7*(18).

Maturana, H. R. (1978). *Biology of language: The epistemology of reality.* In G. Miller & E. Lenneberg (Eds.), *Psychology and Biology of Language and Thought* (pp. 28–62). New York: Academic Press.

Maturana, H. & Varela, F. (1980). *Autopoiesis and cognition: The realization of the living.* Dordrecht, Holland: Reidel.

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4,* 135–183.

Nilsson, N. J. (1984). Shakey the robot. *Technical Note 323,* SRI AI Center, Menlo Park, CA.

Pattee, H. H. (2001). The physics of symbols: Bridging the epistemic cut. *Biosystems, 60,* 5–21.

Pentland, A. (2005). Socially aware computation and communication. *IEEE Computer, 38*(3), 33–40.

Port, R. & Leary, A. (2005). Against formal phonology. *Language, 81*(4), 927–964.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought.* Cambridge, UK: Cambridge University Press.

Ross, D. (2006). The economics and evolution of selves. *Journal of Cognitive Systems Research, 7,* 246–258.

Schank, R. G. & Ableson, R. P. (1977). *Scripts, goals, plans and understanding.* Hillsdale, NJ: Erlbaum.

Seabra Lopes, L. & Chauhan, A. (2007). How many words can my robot learn? An approach and experiments with one-class learning. *Interaction Studies, 8*(1), PPP–PPP.

Thelen, E. & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Tomasello, M. (2003). *Constructing a language.* Cambridge, MA: Harvard.

Ullman, S. (1980). Against direct perception. *Behavioral and Brain Sciences, 3,* 373–415.

Viger, C. (2007). The acquired language of thought hypothesis: A theory of symbol grounding. *Interaction Studies, 8*(1), PPP–PPP.

Vogt, P. & Divina, F. (2007). Social symbol grounding and language evolution. *Interaction Studies, 8*(1), PPP–PPP.

Worgan, S. F. & Damper, R. I. (2007). Grounding symbols in the physics of speech communication. *Interaction Studies, (8)*1, PPP–PPP.

Vygotsky, L. (1986). *Thought and language.* London: MIT Press.

Wason, P. C. (1981). Understanding the limits of formal thinking. In H. Parret and J. Bouveresse (Eds.), *Meaning and Understanding.* Berlin: Walther de Gruyter.

Wittgenstein, L. (1953). *Philosophical investigations.* Oxford: Blackwell.

### *Author's address*

Karl F. MacDorman
Indiana University
School of Informatics
IT 487, 535 West Michigan Street
Indianapolis, Indiana 46202 USA

http://www.macdorman.com

### *About the author*

**Karl F. MacDorman** is Associate Professor of Informatics at Indiana University and Adjunct Professor of Science and Technology at Purdue University. He focuses on human–robot interaction and the symbol grounding problem. He has co-organised the workshop Toward Social Mechanisms of Android Science at CogSci 2005 and CogSci/ICCS 2006, the workshop Views of the Uncanny Valley at IEEE Humanoids 2005, and the special session Psychological Benchmarks of Human–Robot Interaction at IEEE Ro-Man 2006 and has edited special issues on these topics for *Connection Science* and *Interaction Studies.* He has published extensively in robotics, machine learning, and cognitive science.