



<http://www.diva-portal.org>

## Postprint

This is the accepted version of a chapter published in *Meaning, Mind and Communication: Explorations in Cognitive Semiotics*.

Citation for the original published chapter:

Parthemore, J. (2016)

A cognitive semiotic perspective on the nature and limitations of concepts and conceptual frameworks.

In: Jordan Zlatev, Göran Sonesson, Piotr Konderak (ed.), *Meaning, Mind and Communication: Explorations in Cognitive Semiotics* (pp. 47-68). Frankfurt am Main: Peter Lang Publishing Group

N.B. When citing this work, cite the original published chapter.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-13249>

## Chapter 3

### A Cognitive Semiotic Perspective on the Nature and Limitations of Concepts and Conceptual Frameworks

Joel Parthemore

#### 1. Introduction: From cognitive science to cognitive semiotics

What philosophy of mind calls *theories of concepts* (for a representative listing of theories, see Section 3.2), cognitive science has long described as *knowledge representation* (Section 3.4). Regardless of the choice of terminology, the concern is with addressing how it is – and what precisely it means to say – that the thought patterns of non-infant human beings (if not others) are systematic and productive: systematic in that the same ideas can be applied in essentially the same way in each new context the agent encounters; productive in that a finite number of these ideas can be combined into unboundedly many complex structures that can, in the human case at least, be expressed in language: e.g., the notion of a *politically left-leaning, possibly Asperger's, science-fiction reading, computer-game playing, bicycle riding American philosopher of mind with a taste for hot chili peppers*, a notion I came up with just now. Indeed, cognitive science – whose roots are often traced back to the Dartmouth conference of 1957 – has been seen by many as preoccupied, since its foundation, with what it means to have a (human) mind – presumably something *more* than a pure stimulus/response system – and whether and in what sense that mind could be said to be “computable”.

Cognitive science, as traditionally practiced, has often and justly been held to criticism (Froese 2007, 2011, 2012) and even scorn (Dreyfus 1972, 1992) and ridicule. Consider this from Anthony Chemero (2011: 6):

Empirical propositions about the number of planets or about the history of cilia, it is typically thought, are not to be ruled out by logic or definition. Somehow, though, this attitude has not made its way into cognitive science, where conceptual arguments against empirical claims are very common. Indeed, one could argue that the field was founded on such an argument.

In offering his own critique, Jordan Zlatev (2012: 3) writes of cognitive science that it has often “from its onset in the 1950s adopted an explicitly physicalist (computational and/or

neuroscientific) take on mind, connecting to the humanities quite selectively, and above all to philosophy of mind with a distinctly reductionist bent”.

Nevertheless, cognitive science in general and the mind-as-computer metaphor in particular have a proud history of more than sixty years of conceptualizing and re-conceptualizing the very structure of conceptually structured thought itself: as clearly demonstrated among human beings, as arguably found in varying degrees among a range of other species, as much more controversially attributed either to existing or potential future artefacts. Together, they have provided a distinctive mirror for the human species to hold up to itself and contemplate – in an inevitably self-referential-paradox threatening fashion – the nature of its cognitive nature, along with that of other agents who share elements of what it means to be cognitively human.

Cognitive science has been criticized by Tom Froese and others for all too often implying, if not actively embracing, some form of mind-versus-matter Cartesian dualism.<sup>1</sup> In case one thinks that Froese has a straw man in mind, Owen Holland – though not a substance dualist – in describing his Chronos/Simnos architecture (see e.g. 2007) openly embraces the sort of “Cartesian theatre” homuncular model that Daniel Dennett sees as intimately tied up with Cartesian dualism.

That said, cognitive science has often if not generally been inclined toward a more pragmatic attitude, as exemplified so well by the work from the 1970s onward at the University of Sussex, UK, of such giants as Margaret Boden, with her hugely influential *Artificial intelligence and natural man* (1977), which shaped many an aspiring cognitive science student; and Aaron Sloman, with his perhaps equally influential *The computer revolution in philosophy* (1978). Sloman I see as representing some of the best in what has come to be known (often disparagingly!) as GOFAI, short for *good old-fashioned AI* (Boden 2006, ch. 10), with his focus on AI as cognitive modeling (as opposed to, say, consciousness creation: see e.g. Sloman, 1985) and his programmer’s/hacker’s mentality of trying things out just to see what they do. Both Boden and Sloman may be seen as proponents of a pragmatic kind of functionalism that assumed by no means a complete independence between surface behaviour and underlying mechanisms – which would imply, if not require, a form of Cartesian dualism – but only a *relative* independence. In particular, theirs was, and is, a form of functionalism that emerged in response to the worst excesses of the earlier behaviourist movement, with its positivist-inspired

---

<sup>1</sup> Substance dualists are notoriously thin on the ground, but physicist Henry Stapp (2007) is one, though his “quantum” interactive substance dualism is meant to avoid the problems of earlier such accounts.

disregard for anything that could not directly be measured. This functionalism emerging out of the 1970s and '80s stressed that, if one were to have any hopes of getting one's cognitive models right and explaining all the behaviour that *could* be directly perceived, then one had to make the right assumptions about all the underlying cognitive processes that could *not* be so perceived.

For all their accomplishments, Boden and Sloman – and many others like them – stand on the shoulders of another giant, whose article in *Mind*, under the unassuming name of “Computing machinery and intelligence” (Turing 1950), aimed straight for the heart of one of the most basic of questions: what does it mean to be a thinking (conceptual) being? Turing was one of the first to see the powerful metaphorical connections between the new digital computers – remembering, of course, that the word “computer” originally applied to *persons* who performed calculations – and the conceptual structure of human thought, raising important questions about how far such thought could be reduced to some amount of computation, some set of algorithms, and where any such explanation must forcibly stop. Although his Imitation Game, introduced in that paper, has long been taken as providing something like a gold standard for establishing “true” intelligence (e.g. Whitby 1996), Blay Whitby and I have argued (Parthemore and Whitby 2014) that Turing never intended it as an operational test for intelligence but rather as a thought experiment to get his readers to think *about* thinking: where does simply going through some set of predefined motions end, and flexible conceptual thought begin? What is it that sets a thinking agent apart from an automaton?

One further name bears mention with relation to the heady days of cognitive science research in the 1970s and '80s: John Searle, with his *Chinese room* thought experiment (Searle 1980): a fascinating counterpoint to Turing's Imitation Game. The Chinese room can and has been taken as a lecture on the sin of disregarding embodiment and its role in cognition (e.g. Preston and Bishop 2002) – a concern I enthusiastically endorse. Nevertheless, I object to the unstated, yet critical, intuition/presupposition that Searle quietly slips in: namely that, “if one looks inside the Chinese room – or, by analogy, inside the skull – sees the operations there, and ‘knows’ that those operations could not even in principle produce intelligence/consciousness, then there is no intelligence/consciousness *regardless of any observable behavior or indeed any measurable differences whatsoever*” (Parthemore and Whitby 2013: 113). Consider the ethical minefield that this opens: if one “knows” that the agent is “really” an automaton, then that might seem to justify whatever one might do to it in terms of torture, abuse, and so on:

after all, one “knows” that it feels nothing, despite all its perhaps frantic claims to the contrary.<sup>2</sup> For Searle, there are no possible grey areas between being an automaton and a thinking agent, but rather (as for Descartes, in his own way) an absolute divide.

I am as convinced as Searle that intelligence and consciousness cannot be achieved by brute-force programming. Nevertheless, what Searle would place beyond possibility of doubt I prefer to leave open in principle to empirical discovery – in the end, my empiricist tendencies trump my rationalist ones. Like Sloman, I want to try things out and see what happens; so the theory of concepts I am developing (see Section 3) has been accompanied, from the beginning, by a software program as a direct translation of that theory: one that shows both where the theory works and where it still has gaps.

Logical fiats – such as Searle appears to offer – should always be viewed with suspicion. For all of their errors and omissions in other regards, the behaviourists had *this* much right: all one can go on, in the end, is what one observes.

Echoing Whitby’s (2003) sentiments, the take-home message I wish to deliver is that, if cognitive science was never that day or week or month or year away from achieving “true” AI – as its more naive or unprincipled advocates declared – it was never quite so blind to situatedness or embodiment or neuroscience, nor so impoverished of accomplishments, as many contemporary researchers have seen fit, for their own rhetorical purposes, to portray it.<sup>3</sup> Nevertheless, like most if not all intellectual endeavours, cognitive science faces the need periodically to re-invent itself – in light of the present age, in keeping with contemporary insights and discoveries.

*Cognitive semiotics* – characterized by Zlatev (2012: 2) as “an emerging field dedicated to the ‘transdisciplinary study of meaning’, involving above all researchers from semiotics, linguistics, developmental and comparative psychology and philosophy” – offers one tantalizingly encouraging opportunity to do so. Cognitive science, he complains, has often been

---

<sup>2</sup> Precisely this territory is explored in the 2015 film *Ex Machina*.

<sup>3</sup> Again, if one thinks that they are erecting a straw man, one need only look at the (growing) field of *artificial general intelligence*, which unabashedly attempts to revive an explicit-rule-based GOFAI-style approach. To be fair to AGM advocates, including Claes Strannegård, whom I cite below, they would certainly not describe themselves as “blind”; they would simply say that “true” AI can be achieved without necessarily needing to address one or more of these issues. See (Nizamani et al. 2015; Strannegård et al 2013) to get a good feel for the general AGM style.

too explicitly reductionist and physicalist; “[cognitive semiotics] is... considerably more pluralist in its ontological and methodological commitments” (ibid: 3).

For all I see as the enduring strengths of Turing’s insights, what is perhaps most striking to me about the mind-as-computer metaphor is how badly so many of both its fiercest advocates and opponents seem to misunderstand the nature of minds *and* computers. Computers do *not* simply “do as they are told” – even as existing models lack what could be called, in anything but the most metaphorical of terms, intelligence, never mind free will. Despite what Roger Penrose (1994: 66) claims to the contrary, they are *not* equivalent to Turing machines – which are strictly idealized mathematical entities. The problem is *not* that computers are not embodied, and embedded in an environment with which they richly interact: they are. The problem is rather that they are embodied *in the wrong way*, not least because of the seemingly unavoidable dependence of cognition on life – even if that life is something different than traditionally biologically conceived (Zlatev 2001, 2002, 2009a).

Meanwhile, for all I disagree with her on other matters, Patricia Churchland is right to observe – *contra* Penrose – that it is unclear whether *anything* in the universe is not, in principle, algorithmically describable or “computable”, in the very broad way of computability Penrose has in mind (Grush and Churchland 1995: 190). That does not mean that an algorithmic description is always desirable, or even possible. After all, some things may, and probably *do*, ultimately outstrip the capabilities of human cognition – including, almost certainly, its capacity to understand itself. Here again cognitive semiotics, with its openness to methodological and theoretical pluralism, shows its strengths, in allowing the possibility that there may be (and probably is) no one single “correct” way to look at the mind, and that the mind-as-computer metaphor, properly understood, has its role to play. The problem is not the metaphor; the problem, again, is the need for periodic renewal. To echo a concern from elsewhere in cognitive semiotics (see below), understanding how the human mind differs from a digital computer – and the gulf is wide indeed! – requires first having an understanding of those things we have in common.

In Section 2, I offer my interpretation of the field of cognitive semiotics. In Section 3, I discuss the focus of my own research interests: namely, the nature of conceptually structured thought. Section 4 offers a cognitive semiotic perspective on concepts, looking at how cognitive semiotics improves over traditional cognitive science in offering insights into the nature of concepts and conceptual frameworks. Section 5 applies these ideas to metaphor theory. Throughout there will be echoes of both the relationship between cognitive science and

cognitive semiotics, and the enduring power of the mind-as-computer metaphor. I close with some thoughts about the way forward both for theories of concepts and for cognitive semiotics.

## 2. The emergence of cognitive semiotics

[Cognitive semiotics] can be seen as called for by historical needs... the need to unify or at least to “defragment” our world-views, the need to come to terms with increasingly higher levels of dynamism and complexity, the need to understand better – and thus deal with – the dialectical relationship between individual freedom (autonomy) and collective dependence (sociality), etc. In other words, if Cognitive Semiotics did not exist, we would need to invent it (Zlatev 2012: 18-19).

Zlatev (2012, 2015) describes how cognitive semiotics emerged as a distinctive field in the 1990s, first in Denmark and then in the US and Sweden. Its focus was, and remains, on the study of meaning and meaning making.<sup>4</sup> In the spirit of enactivism (Thompson 2007; Maturana and Varela 1992; Varela et al. 1991) with its focus on agent/environment interaction and the irreducibility of that interaction, “a basic... tenet is that meaning is not ‘inside’ brains, minds, groups, or communities but is a result of processes of self/other/world interaction” (Zlatev 2012: 17). Meaning is fundamentally perceptually grounded. Thus, one of the leading voices of cognitive semiotics, Göran Sonesson “has consistently argued for the primacy of perceptual meaning over other kinds of meaning – including signs” (Zlatev 2012: 6).

Signs and sign use are the most obvious “outward” indicators of meaning. Cognitive semiotics finds some significant part of its origins in semiotics. That said, “it is *not* to be seen as a branch of the overall field of semiotics, defined either in terms of ‘domain’ (in the manner of e.g. biosemiotics, semiotics of culture or social semiotics), or ‘modality’ (e.g. visual semiotics, text semiotics)” (Zlatev 2012: 2). In keeping with this relationship, cognitive semiotics tends, when it talks about signs, to prefer a far narrower definition than offered by e.g. Charles Sanders Peirce, for whom a sign may be taken to include nearly *anything* to do with meaning (see e.g. Peirce 1981, in stark contrast to Sonesson 2007) – but a broader definition than one that restricts signs to language or human-language-like communication.

For myself, I prefer to take signs as conceptually mediated attempts at communication that are at least semi-conventionalized, of which language is a subset; the attended target could be another agent, or it could be oneself (though the former is presumably primary). Signs and

---

<sup>4</sup> Or sense making: see e.g. (Thompson and Stapleton 2009; De Jaegher and Di Paolo 2007).

sign use presuppose communication and social context. Like concepts (see Section 3), signs have both expression and content, each of which may be distinguished from the other. Unlike concepts, expression and content must be distinguishable – at least normally – not just to an “outside” observer but to the sign-using agent herself. On such a view, concepts are prior to signs and sign use ontogenetically and phylogenetically, while signs and sign use are prior to language.

Zlatev (2012, 2015) describes cognitive semiotics as intrinsically interdisciplinary – or, as he prefers, *transdisciplinary*. Cutting across rigidly delineated academic subdomains is, of course, hardly new to cognitive semiotics. What *is* new is the particular mix of disciplines it brings and, in particular, the way it takes lessons from semiotics (regarding meaning), phenomenological philosophy (regarding experience), and linguistics (regarding language) to re-invigorate discussions that have been going on for half a century in cognitive science. Of course, interdisciplinarity is a two-edged blade: focus too narrowly on one’s specialization, and one risks facing endless terminological turf wars when meeting researchers who use the terms differently; focus too broadly, and one risks losing oneself in terminological and conceptual vagueness. Different disciplines use terminology differently for good reasons: because their needs for that terminology, though overlapping, are also different. Nevertheless, a certain disregard if not irreverence for traditional academic boundaries – as cognitive semiotics clearly shows – is hard to see as anything but a good thing; as is its refusal to be drawn either toward a theoretically strong but empirically weak perspective (all too common in philosophy), or an empirically strong but theoretically weak one (all too common in empirical science; see Zlatev 2012: 14).

Finally, taking both an onto- and phylogenetic perspective, cognitive semiotics attempts to discover what is uniquely human through a better understanding of what is not,<sup>5</sup> with the understanding that what *appears* to be key, at first blush, may not be. So far, language – in the sense of a rich integration of syntax, semantics, and pragmatics and not merely innovative communication – seems to be. On the other hand, concepts and conceptually structured thought arguably are not (see e.g. Parthemore 2011a, 2013, 2014).

Cognitive semiotics is not an attempt to replace cognitive science. Neither is it, as Zlatev (2012: 3) considers and rejects, “a new and fancier name” for it. It is rather an attempt to shine

---

<sup>5</sup> Compare: “...while CS [cognitive semiotics] practitioners indeed focus on what is specific about human forms of meaning-making, there is widespread agreement that this can only be properly understood in a comparative and evolutionary framework” (Zlatev 2012: 2).



new light on old phenomena: whether that be the origins of proto-language and language or of consciousness, or the limitations of AI and so-called *machine consciousness* – or the nature of conceptually structured thought.

### 3. Concepts and theories of concepts

#### 3.1 Concepts: Attempt at a definition

Concepts are the framework underlying not just language but unspoken thought. They may roughly be taken equally either as units of structured thought or as the ability/-ies to possess and employ structured thought – I take the two formulations as equivalent; such that that thought is (a) systematic... (b) productive... (c) compositional... (d) intentional *per* Brentano ... (e) re-identifiable... (f) “spontaneous” *per* Kant’s terminology... and... (g) subject to revision... (Parthemore 2014: 193-194).

What this definition amounts to is that a concept of  $x$  is one’s structured understanding of some concrete or abstract entity  $x$ , such that one might, but need not, be able to explain it even to oneself, let alone anyone else. After all, many of the skilled *basic activities* we engage in we feel like we “just do”, without any need to think about them (Felix 2015: Ch. 3). This claim – that concepts, in effect, bridge Gilbert Ryle’s (1949) *knowledge how / knowledge that* divide – will be controversial in certain circles; as will the claim that concepts must be relatively stable but not *too* stable: that they are not just intrinsically *open* to change but indeed to continuous, if incremental, revision.<sup>6</sup> Concepts do not stand still. They must be stable to be reusable across unboundedly many contexts but not *too stable* so as to be unable to adapt to each new context – in some way, large or small, different in a thousand details from any context that preceded it and any that follow.

---

<sup>6</sup> Jerry Fodor would oppose it outright: on his Informational Atomism account (Fodor 2008), it is essential to the nature of concepts that they *cannot* change. Others would say merely that they *can* change.

The other properties I list are, by contrast, all but universally accepted. I have already explained systematicity and productivity. Compositionality – the idea that concepts can be added together (and at least some concepts can be taken apart) – falls out directly; as does re-identifiability: the ability to describe the same thing as the same thing – from many different perspectives, in unboundedly many contexts, by various sensory modalities.<sup>7</sup> Intentionality<sup>8</sup> is the deceptively simple idea that concepts are always concepts *of*: they always have some target (indeed, for someone).

All of this, however, is to describe concepts as if they were carefully pinned out lepidoptera. What are concepts *in use*?

Concepts simplify (and, by simplifying, distort) in pursuit of understanding, reducing the complexity of the world to a level we can grasp. They allow us to approach the world in a distinctively flexible way at the price of removing us *from* that world – in favour of a conceptual model of it.<sup>9</sup> That flexibility of response is the hallmark of both conceptually structured thought and consciousness. Indeed, conceptual agency and consciousness may be seen as two sides of a single coin (Parthemore and Whitby 2014), where the appropriate attribution of the one is the appropriate attribution of the other. On my reading, such a link is assumed by most contemporary theories of consciousness – Information Integration Theory (Tononi 2008; Balduzzi and Tononi 2008), Global Workspace Theory (Baars 1988, 1996), Dual Aspect Theory (Chalmers 1997), to name just three – but rarely if ever made explicit, an omission I wish to correct.

In keeping with arguments made by Douglas Hofstadter (2000 [1979]) with respect to Gödel’s Incompleteness Theorem (see also Parthemore 2011: 192-195), concepts logically entail their own limitations. What this means is – *pace* Penrose (1994) – that any given conceptual framework cannot be complete and consistent at the same time, as a conceptually well-structured argument, derived from Gödel’s two proofs, can be made to show. Otherwise,

---

<sup>7</sup> A formulation I owe to Ruth Millikan (personal communication).

<sup>8</sup> *Per* Brentano. Philosophers also speak of intentionality in another way, as that which is willfully motivated: e.g. Felix (2008), following Davidson (2001), uses it in this sense.

<sup>9</sup> By saying this, I reveal not just my phenomenological leanings but my antirealist and Kantian metaphysics: concepts *are* interpretation; the world is essentially always revealed to us interpreted, because we cannot set our conceptual frameworks aside. Neither can we thoughtfully reconstruct the uninterpreted world. There is no “literal” meaning that can simply be taken for granted (see Section 5).

the very expressive power of conceptual frameworks allows the formation of paradox-inducing self-referential conceptual structures that threaten to bring the entire framework down. Indeed, as Hofstadter suggests, one might imagine – in a manner reminiscent of the ancient Greek Pyrrhonists (Lind 2013) – deliberately exploiting the paradoxes to achieve just that.

Putting this in another way, concepts portray the world as being one way and not another, with no conceptual perspective privileged above all others: a conceptual perspective is necessarily a limited one. That most pointedly applies, in paradox-threatening fashion, to a conceptual view on concepts themselves. For a good visual equivalent, consider what happens if you train a camera on its own playback monitor. If the focus is a bit off, one gets an amusing regress (see Figure 1); if it is straight on, one gets an attractive but uninterpretable abstract pattern.



**Figure 1.** "Video feedback at last night's NYU IMPACT performance": public-domain image downloaded from Wikimedia Commons (<http://commons.wikimedia.org>).

At least for human conceptual agents (simpler conceptual agents may be less encumbered), concepts have a habit of getting very abstract very quickly, particularly when we stop to contemplate ourselves contemplating them, as opposed to merely getting on with using them non-reflectively. Consider the concept of *grasp*, highlighted by Vittorio Gallese and George Lakoff (2005) in making their claim that all that *any* concept amounts to is nothing more than a specific sensorimotor engagement, albeit with parts of that engagement typically suppressed. Their claim is oddly reminiscent of George Berkeley (1999 [1710]), with his claim that no one

has a general concept of “triangle” that is not the image of a specific triangle, albeit with parts of that image ignored.

Their choice of such a seemingly concrete concept is oddly telling, since their account is meant to apply to *all* concepts – even the most abstract. The problem is that even the concept of *grasp* is not so clearly an “actual” sensorimotor grasping as Gallese and Lakoff would have one believe. *Pace* their account, the concept of *grasp* is at least as unlike a specific instance of grasping as it is like. The very power of the concept – the power of concepts in general – is the way it abstracts away from any specific instance of application at the same time that it applies back to every conceivable application: both the cases of physical grasping and those – such as “eureka” moments – where the grasping is noticeably metaphorical/abstract.

### 3.2 Theories of concepts

Within philosophy of mind, theories of concepts fall within a subdomain, active at least since the late 1990s, whose leading figures include the following.

- Jerry Fodor, with his *Informational Atomism* account (1998), whereby most concepts are non-decomposable physical-symbol-based atoms that track their referents in the world in law-like fashion: e.g., a concept of *gold* is a concept of *gold* because it tracks all and only gold (and not, say, philosophers),
- Jesse Prinz, with his *Proxytypes Theory* (2004), which he describes as “informational atomism without the atomism” (ibid: 164), and
- Peter Gärdenfors with his *Conceptual Spaces Theory* (CST) (2004) – based, like Proxytypes Theory, on Eleanor Rosch’s work on *prototypes* (1975, 1999); as well as, perhaps,
- Ruth Millikan, who offers no theory of concepts as such, but sees concepts less as representations (say) than *abilities* to form representations<sup>10</sup> and stresses how accounting for them requires a teleological account (Millikan 1998, 2010), and
- Edouard Machery, who argues (2009) that concepts do not constitute a proper class of things at all, but several disjunct sets: one set belonging to psychology, another to philosophy, another to linguistics, etc.

My own *Unified Conceptual Space Theory* (UCST) (Parthemore 2011b, 2013, 2015; Parthemore and Morse 2010; see below) is largely based on and attempts to extend Gärdenfors’

---

<sup>10</sup> A formulation I owe to Millikan, via personal communication.

work in a more algorithmically well-defined and empirically testable direction, offering a precise algorithmic “recipe” for the creation, evolution, destruction, and application of concepts – both at the level of the individual and that of the group or society. It comes with the aforementioned software program, intended to be used as the basis for a series of experiments designed, with a psychologist, to test the underlying theory (Parthemore 2015).

### 3.3 The Unified Conceptual Space Theory (UCST)

UCST attempts to fill in the gaps in the self-avowedly scaffold-like structure that is CST. In the process, it makes explicit the commitment to enactive philosophy that CST leaves implicit; talks of representations<sup>11</sup> in a much narrower, more precise way, borrowed from Inman Harvey (1992), according to which *agent A intentionally (with wilfully motivation) uses B to stand in place of C for herself or another agent D*; and, most importantly, provides a framework within which all of an agent’s (or, on another level, society’s) many conceptual spaces, as discussed by CST, can come together in a single, unified *space of spaces*.

UCST makes the following empirically testable claims (see Parthemore 2015):

1. All concepts can be assigned to one of three sub-categories (*objects, events, and properties*) that derive from three innate proto-concepts (*proto-objects, proto-events, and proto-properties*): so-called because they fail to fulfil all the properties of “true” concepts (in particular, they are too few in number to be, of themselves, productive; and, being innate,<sup>12</sup> they are not under the agent’s endogenous control). These sub-categories align, more or less, with the English lexical categories of noun/pronoun, verb, and adjective/adverb.
2. All concepts can be oriented with respect to one of three dimensions that define the unified space and that represent the *integral dimensions* (see Gärdenfors 2004: 24-26) in common to all concepts: namely, an *axis of generalization* from most specific to most abstract, reprising the familiar “ISA” hierarchy whereby e.g. a cat is a mammal is an animal is an organism, etc.; an *axis of abstraction* from most concrete and seemingly non-conceptual/physical to most abstract and straightforwardly (higher-

---

<sup>11</sup> Needless to say, this disallows talk of e.g. *neural representations*. Meanwhile, many would talk of *mental representations* as something ontologically distinct from so-called *external representations*, like paintings. UCST allows that such a distinction is, at best, a useful conceptual distinction and not a prior ontological one. What makes *anything* a representation is the perspective that an agent intentionally takes toward it.

<sup>12</sup> In this way, UCST endorses a very modest nativism, in stark contrast to the radical nativism commonly associated with the early Fodor, whereby all – or nearly all – of our concepts are innate.

order) conceptual/mental, where all points along the axis represent one and the same target at different levels of abstractness; and, finally, an *axis of alternatives*, along which one finds all the possible variations of a concept at the same level of generality/specificity and concreteness/abstractness, derived by varying the values of any one or more of the integral dimensions defining the concept most specifically (e.g., *hue*, *saturation*, and *brightness* in the case of colour).

3. In addition to these proximal connections to “neighbouring” concepts within the unified space, concepts relate to distal (“non-neighbouring”) concepts in three possible ways:
  - a. Some concepts (more concrete object and event concepts, but not abstract objects or events, or properties in general) have one or more necessary subcomponents of the same broad category (the subcomponents of object concepts are always object concepts, that of event concepts always event concepts).
  - b. All concepts are defined by one or more integral dimensions.
  - c. All concepts exist within a context of other concepts with which they are more or less routinely associated (co-occurring).

A useful metaphor is children’s building blocks, but a special kind where blocks affect not only directly adjacent blocks but ones that may be considerably removed in the pile.

### 3.4 Knowledge representation

Within cognitive science, the same discussions covered in philosophy of mind under theories of concepts have traditionally fallen under the rubric of *knowledge representation* (KR), with a focus on propositionally structured (or structurable) knowledge such as might be expressed in the following statement from the programming language Prolog, which may be read as saying that two persons are first cousins if each has a parent such that those parents are siblings:

first-cousin(X,Y) :- parent(R,X), parent(S,Y), sibling(R,S).

What one might call the “KR view on concepts” is problematic for UCST because, like CST, UCST takes concepts to be beholden neither to *knowledge that*, as an “intellectualist” like Fodor or Jason Stanley (2013; see also Felix 2015: chs. 4-5) would have it; nor *knowledge how*. Concepts sit in the middle, being both *knowledge how* and *knowledge that* and, at the same time, neither. They are neither particularly propositionally amenable (as *knowledge that* is traditionally understood) nor propositionally averse (as *knowledge how* is portrayed by all but the intellectualists).

Put another way, concepts are neither strictly representational (as the KR view suggests) nor strictly non-representational and associational (as Rodney Brooks might have it: see e.g. Brooks, 1991a, b). Again, concepts sit in the middle: looking more like simple associations when pushed the one direction, more like iconic and symbolic representations when pushed in the other.

It follows, *contra* the KR view, that concepts are neither intrinsically linguistic nor – for linguistic agents like human beings – properly separable from language. Language transforms and extends our conceptual abilities, not only facilitating the sharing of concepts but allowing existing concepts to become more abstract and making possible the contemplation of abstract concepts that probably could not arise without language (Parthemore 2014: 197-198). An agent who lacks language will have concepts more toward the *knowledge how*, more toward the non-representational and associational, more toward the first- as opposed to higher-order, more toward the concrete than the abstract – with all of these positioned along a continuum.

Most worrisome of all, perhaps, the KR view on concepts risks conjuring what might be seen as the worst renderings of the mind-as-computer metaphor: the mind as a physically implemented Turing machine, processing through an explicit set of linearly arranged instructions:<sup>13</sup> the kind of soulless automaton imagined by Descartes for all non-human animals, extended to human beings by T.H. Huxley (1874; see also Kim 2007). Once again, cognitive semiotics offers a contrasting perspective.

#### **4. A cognitive semiotic perspective on concepts**

As suggested in Section 2, cognitive semiotics provides a way to consider how concepts on the one hand and language on the other (or sign-based communication more generally) are both deeply intertwined and pull apart. Such a position is controversial in some circles, as concepts are often taken to equate to lexical concepts – either as a matter of logical stipulation or empirical consequence. Nevertheless, the opening for such a position is provided by the broad view cognitive semiotics is inclined to take toward communication and cognition, and the

---

<sup>13</sup> An anonymous reviewer objected, at this point, that I was addressing an “outdated version” of cognitive science; I hope I have convinced the reader of reasons to think otherwise. AGI aside, I could offer as further evidence of the persistence of what I am calling the “KR view” – just to take three examples – SNePS (Shapiro, 1978; Shapiro and Rapaport 1987), ACT-R (Anderson 1996), and Cyc (Lenat 2006): all deeply “symbolic” systems.

consequent “conviction that language – its nature, evolution and development – cannot be understood outside the context of a more general approach, taking both meaning and mind seriously” (Zlatev 2012: 7). Meanwhile, the arguments from philosophers in favour of “animal concepts” (Newen and Bartels 2007; Allen 1999) and the empirical evidence for complex cognitive capacities – in effect, conceptual abilities – in non-human species (including corvids and delphinids, as well as the great apes; see e.g. Jacobs et al. 2015; Osvath and Sima 2014; Osvath et al. 2014; Osvath and Persson 2013; Osvath and Karvonen 2012; Osvath 2009; Bugnyar 2011; Raby et al., 2007) continue apace.

Figure 2 shows my proposed revision or re-interpretation of Zlatev’s (2009a) *semiotic hierarchy*, with its four successive levels in the organization of meaning: life, consciousness, sign function, and language, with each level building on the previous. Because of what I see as their intimate relationship, I place *concepts* on the same level as *consciousness* and add two additional levels, (intentional inter-agent) *communication* between *consciousness/concepts* and

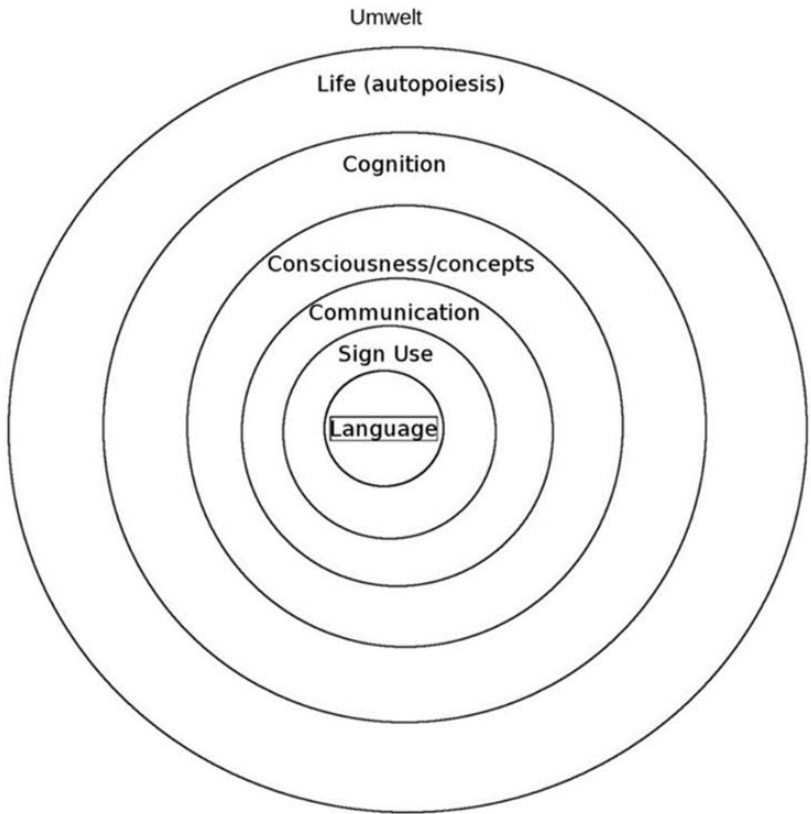


Figure 2: A revised version of the semiotic hierarchy (cf. Zlatev 2009a).



*sign use*,<sup>14</sup> and *cognition* between *life* and *consciousness/concepts*. The context for life itself is, of course, the *umwelt* – which becomes, for the conscious agent, the *lifeworld*.

Drawing on its roots in phenomenology, which takes experience as a fundamental entity – ultimately *explanans* rather than *explanandum* – cognitive semiotics focuses attention on the conceptual distinction (as opposed to prior ontological one) between concepts as we reflect on them and concepts as we possess and employ them (presumably most of the time) non-reflectively. It is the difference between concepts at the level of *pre-reflective self-awareness* and that of *fully reflective self-consciousness* (e.g. Gallagher 2007). Of course, there is no one single school of phenomenology; in particular, there is a contrast to be made between those phenomenologists, following certain readings in Husserl, who would use phenomenological methods in general and the *phenomenological reduction* in particular to pass beyond experience to the world itself, and those who take what might be seen as a more Kantian line, seeing the mind-independent world and the *lifeworld*, the *Körper* and the *Leib*, as ultimately inextricably intertwined: all that one sees, one sees from a perspective, and no perspective is privileged above the others. Regardless, “...the basic idea is to depart from experience itself, and to provide descriptions of the phenomena of the world, including ourselves and others, as true to experience as possible – rather than constructing metaphysical doctrines, following formal procedures, or postulating invisible-to-consciousness causal mechanisms that would somehow ‘produce’ experience” (Zlatev 2012: 15).

Given such an approach, it seems difficult if not impossible both to give full credit to the way we experience our conceptually structured thought, and to allow that the way we experience it may not be – indeed, most likely is not – revealing of its pre-reflective nature. Table 1 gives some idea of how these two views on concepts – reflective and pre-reflective – give rise to a whole series of contrasting perspectives on the nature of concepts.

Through its focus on meaning as dynamics and on meaning making in the context of dynamical systems, cognitive semiotics shows how concepts, though they are concepts to the extent they are relatively stable across contexts, can only function as concepts to the extent that they are able to change as needed to fit each new context: to revise the old proverb, one cannot

---

<sup>14</sup> I intend *communication* to be more-or-less equivalent to the stage Zlatev himself adds in the current, five-stage version of the model, which he terms “culture”, even if the emphasis is slightly different (Sonesson *in press*, Zlatev *in press*). Note that I am using “communication” in a narrow sense; (Zlatev, 2009b) applies the model as a whole to communication in the broad sense, showing which kinds of communication are possible at each level.

fit an unchanging peg into a constantly changing hole; the essentially static nature of concepts is an illusion, and meaning only appears fixed as a necessary convenience (Parthemore 2014).

**Table 1.** Two contrasting yet complementary perspectives on concepts.

Concepts as...	Concepts as...
things we reflect upon	things we possess and employ non-reflectively
intentionally imposed “top down”	activity derived “bottom up”
product of reason (rationalism)	product of discovery (concept empiricism)
objects of perception	means of perceiving objects
consciously accessible	partly or substantially not consciously accessible
<i>knowledge that</i>	<i>knowledge how</i>
intellectual	pre-intellectual
symbolic entities	skilful abilities
“mental” representations	abilities to form representations
sub-propositional components of thought	subconscious components of interaction
abstract and “mental”	concrete and “physical”
abstracted from context	sensitive to context
“internal” to agent	“external” to agent – in the environment
static	dynamic
discrete (individuable)	continuous
easily tied to language	clearly distinct from language
private	public

The question of whether concepts change (Woodfield 1990) and, if so, how intrinsic is that capacity for change (do some change? all? must they change?) remains one of no small debate. Fodor’s concepts cannot change because concepts, on his account, are strictly public entities defined purely by their extension: my concept of gold is exactly the same as the ancient Greeks’ concept of gold, despite our differing understandings of and beliefs about gold, because the thing that *gold* refers to is the same. On a cognitive semiotic perspective, in contrast, beliefs are partly constitutive of concepts and not just (as Fodor and everyone else allows) concepts constitutive of beliefs: i.e., concepts are defined both by their extension and intension.

Cognitive semiotics is so open to not just the possibility but the inescapable reality of conceptual change because of the way it “studies meaning on all levels – from perception to

language, along with the various forms of 'external', cultural representations (theatre, music, pictures, film, etc.) – primarily as dynamic processes rather than static products.... Nearly all [cognitive semiotics] scholars have made the point that viewing meaning in purely static, structural terms is insufficient for understanding the essentially relational, subject-relative, and (often) interpretive nature of semiosis” (Zlatev 2012: 16). Cognitive semiotics views meaning as fundamentally dynamic, because of the way it views not just life but *lifeworld* itself as fundamentally dynamic.

In alignment with enactive philosophy, which takes interaction as a fundamental entity and agent as continuous with (not ultimately separable from) environment, cognitive semiotics shows how concepts impose crisp boundaries on what are, from a conceptual perspective, underlying continuities. In psychological parlance, this is the phenomenon known as *categorical perception*. As Stevan Harnad explains (1990: ix), “categorical perception occurs when the continuous, variable, and confusable stimulation that reaches the sense organs is sorted out by the mind into discrete, distinct, categories whose members somehow come to resemble one another more than they resemble members of other categories”. He continues (1990: 3): “a [categorical perception] effect occurs when... a set of stimuli ranging along a physical continuum is given one label on one side of a category 'boundary' and another label on the other side.... In other words, in CP there is a quantitative discontinuity in discrimination at the category boundaries of a physical continuum...”.

The implication is that, *pace* the natural kinds philosophers, so-called *natural kinds* concepts do not, to borrow a well-worn phrase from Plato's *Phaedrus*, “carve nature at its joints”; if nature has joints at all, if the notion of natural “joints” is even coherent, then these are not the joints that matter conceptually. This is not to say that concepts (for the individual or the society) are free to draw lines where they will: they are not, any more than agents are free to believe whatever they wish about the world; only that even such basic things as our seemingly inescapable tendency to view the world in terms of *objects*, *events*, and *properties* may say more about us than it does about the world: we literally cannot imagine seeing the world any other way.

With its refusal to shy away from paradoxical or even seemingly contradictory perspectives, cognitive semiotics shows how concepts hide the very world they reveal, beneath a veil of interpretation. They both vastly extend cognitive capacities and, at the same time, in fundamental ways, limit them. On the one hand, they facilitate an active choice of response: faced with the “same” circumstances on different occasions, the same agent can choose differently, based on her reasoned consideration of what has gone before and what is yet-to-

come. Her mind is not trapped in the present moment but free for *mental time travel* (see e.g. Suddendorf and Corballis 2007; Osvath and Gärdenfors 2007) into past and future. On the other hand, they rob the agent of a certain essential spontaneity: a fully in-the-moment experience that, perhaps, only strictly non-conceptual agents can have; at the same time that, by presenting the world in one way rather than another, they blind agents to the range of possible alternative perspectives. Conceptual abilities are neither a straightforwardly good nor bad thing; they *are* something that set us very much apart, though, from species that lack those abilities.

Finally, through its theoretical and methodological pluralism, cognitive semiotics reveals how concepts are not one unified phenomenon. No single theory will ever capture conceptual abilities and frameworks in all the ways one might be interested in. Cognitive semiotics is in general open to the possibility that there is no one, single, “correct” perspective on many if not most matters of substance, including the endless tedious debates over representationalism and whether concepts are best seen as representations (as Fodor or Prinz would have it) or abilities (as Evans (1982) is commonly read, as Gottlieb Frege (1951) doubtless intends, and as Alva Noë (2004) would have it; see also Parthemore 2011a: 20-30). Given its rejection – with the enactivists and phenomenologists – of unconscious representation but its openness to a *kind* of (conscious or reflectively self-conscious) representationalism, cognitive semiotics is amenable to the possibility that concepts both *are* (when we look at them) and are not (when we simply *use* them) representations (see again Table 1). That is, cognitive semiotics advocates a *measured* representationalism/anti-representationalism and rejects the dogmatic anti-representationalism that many discern in radical enactivism (see Zlatev 2012: 12).

Where does all this conceptual revisionism leave us in terms of the mind-as-computer metaphor? Just as digital computers were named after their human predecessors, because of a certain striking resemblance: in particular, the computers’ capacity to perform calculations that, previously, only a human computer could do; so, too, may the human mind be seen to reflect the operations of a digital computer – *in some ways* – albeit a massively parallel one, with a storage and computing capacity that, in certain key respects, continues to put computers to shame. Moreover, certain living creatures, including at minimum human beings, are conceptual agents – something that, given all available evidence, no existing artefact can lay claim to.

At the same time, the take-home message that cognitive semiotics invites us to draw from Brooks among others is that much of our vaunted intellectual capacities – supposedly so flexible, so subject to our free will, so demonstrative of our superiority over other species – are surprisingly mechanical, straightforwardly algorithmic, and utterly dependent on the environment being a certain way. Much of what we think of as conceptual thought is not,

because it is not under our endogenous control. Perhaps we shy away from the mind-as-computer metaphor because we see too much of the “mindless” computer in us.

### **5. A cognitive semiotic perspective on concepts: The case of metaphor**

Before I close, I would like to offer a more specific and concrete example of just how cognitive semiotics can re-shape our thinking about concepts, by looking at how it can help us re-conceptualize not just this one metaphor but metaphor in general – not as a strictly linguistic phenomenon but, in keeping with Conceptual Metaphor Theory [CMT] (Lakoff and Johnson 2008 [1980]; see also Fusaroli and Morgnani 2013) first and foremost a conceptual one, underlying and making possible linguistic metaphor.<sup>15</sup>

CMT has undoubtedly contributed to the study of language and cognition with its recognition that what is metaphorical in language is first metaphorical in thought: again, language and systematically structured thought, though deeply intertwined, pull apart. Nevertheless, CMT remains burdened with conceptual baggage from its predecessors: most notably a tendency toward a strict ontological (as opposed to pragmatic conceptual) distinction between literal and metaphorical and a consequent naive assumption that literal meaning can more or less be taken for granted, while metaphorical meaning requires explanation. This leads to many a tortuous debate over whether a given construction is or is not metaphorical and whether, if it is, it qualifies as “dead” metaphor.

In keeping with CMT, and in alignment with the discussion above, I hold that metaphor is not a strictly linguistic phenomenon. In contrast to CMT, and again following from the earlier discussion, I wish to suggest that metaphor is not definable in opposition to literal meaning, at least without a full explanation for what counts as literal meaning, which Allwood (1981) argues to require as much explanation as metaphor; that it is not clearly demarcated from non-metaphor; and that it is consequently not meaningfully separable into “living” and “dead”. Metaphor is not some type of natural kind, nor indeed any other reifiable entity. It is not something one possesses a stable collection of so much as something one actively and continuously engages with and helps to create.

What, then, is the relation, between literal meaning and metaphor? The literal/metaphorical distinction – often useful, perhaps necessary – is a conceptual rather than a prior ontological one. Alternatively and, I believe, equivalently, metaphor may be defined as a special case of contextual meaning determination, based on meaning potentials (Allwood

---

<sup>15</sup> This is based on work I am currently doing with Jens Allwood at the University of Gothenburg, Sweden.

2003). From a UCST-inspired view, essentially *all* conceptual meaning involves the mapping across domains meant to be characteristic of metaphor, precisely because conceptual experience and non-conceptual world do not cleanly pull apart (Parthemore 2011b). Consequently, metaphor needs to be defined in a different way, one that does not rely on the presupposed nature of the literal.

What distinguishes metaphors is their tendency to map a more abstract domain onto a more concrete one and so give the abstract a *faux* concreteness. In place of a clear, crisp literal/metaphorical divide, one finds a continuum of meaning from primary to secondary, tertiary, and altogether novel meanings/usages, with no sharp lines at any point. Where one draws the lines at any given time is dependent on context: the lines are pragmatic, not absolute. The conclusion is not that all meaning is metaphorical – as some might misread the claim – but that some traces of the metaphorical remain in even the most standardized of meanings/usages and *vice versa*.

A cognitive semiotics perspective on metaphor reveals metaphor as a special kind of sign, according to the definition of “sign” I offered in Section 2. That which is paradigmatically metaphorical is that which – with varying degrees of explicitness – calls attention to itself as metaphor, presenting itself not as an error in usage (either slip or failure in competence) but deliberate act, often taking something abstract and giving it more concrete expression. Linguistic metaphor is generally (though not always) for the benefit of another agent; conceptual metaphor is primarily for one’s own sake – while obviously serving to shape one’s language. With its non-standard use of meaning and its more-or-less deliberate juxtaposition of domains that seem far removed from one another and are not normally juxtaposed, metaphor of whatever kind – conceptual or linguistic, “internal” or “external” – challenges the conceptual agent to ferret out the intended meaning, through some degree of pre-reflective or self-reflective contemplation, thereby experiencing her social and physical environment in new ways.

Such a perspective on metaphor explores metaphor in all its phenomenological richness. Not limiting metaphor to one form or sensory modality, it shows how visual metaphor is as much metaphorical as linguistic or conceptual metaphor (see Figure 3) and opens up such possibilities as tactile, gustatory, and olfactory metaphor, along with various combinations. It finds little use, most of the time, in dividing metaphor between “living” and the “dead” but rather sees metaphor – being a subdomain of conceptual meaning – as deeply and intrinsically dynamic, so that even as that which was living may die, that which was dead may be given new life. Metaphor that ceases to exploit novelty or is no longer able to do so whatsoever – i.e., metaphor that becomes *strictly* conventionalized – ceases to be meaningfully interpretable as

metaphor.<sup>16</sup> Metaphor is not just dynamic but often deeply creative, lending itself to “thinking outside the box”, turning concepts and language, ideas and images, from just another tool into something playful. When one’s view of the world becomes too settled, too mundane, metaphor enters in.



**Figure 3.** “This is your brain on drugs”: centerpiece of a popular and widely memed campaign by the US government as part of its “war on drugs” in the late 1980s and early ’90s. Public-domain image downloaded from Wikimedia Commons (<http://commons.wikimedia.org>).

This perspective on metaphor eschews talk of the literal and sees the seemingly sharp boundaries between metaphorical and non-metaphorical in light of categorical perception. The upshot is that the sharp boundaries are a product of how we use and interpret metaphor and not part of its underlying ontological nature, whatever that might be. Such a move makes questions

---

<sup>16</sup> Consider the English word “haven”, whose Nordic equivalent continues to retain its earlier meaning of “port” or “harbour”. In English, what had been the metaphorical meaning has become the primary meaning, and the earlier primary meaning has been all but completely forgotten.

of what precisely is and is not metaphor highly dependent on context, and arguments over such – whenever some global perspective is assumed – ultimately meaningless.

This perspective on metaphor reveals how metaphor both tremendously extends our cognitive and linguistic possibilities – allowing the juxtaposition of nearly any conceptual domain with any other domain – and, at the same time, limits them. It does so in part by opening the possibility to yet more confused communications and confused thoughts (just because any two domains can be juxtaposed does not mean that they should be: a good metaphor should not be “too” novel but just novel enough, in part by lending the abstract – the most common target of metaphor – a seeming concreteness that it does not genuinely have). The lesson may be that much of what we think of as paradigmatically concrete may be far more abstract than we realize: not least our seemingly most concrete concepts.

Finally, through its theoretical and methodological pluralism, cognitive semiotics shows how metaphor can usefully be considered through its own sets of contrasting perspectives (Table 2).

**Table 2.** Two contrasting yet complementary views on metaphor.

<b>Metaphor as...</b>	<b>Metaphor as...</b>
conventionalized (based on degrees of conventionality)	novel (valued for its novelty)
implicit	explicit
conceptual (“internal”)	linguistic/visual/etc. (“external”)
tied to a certain form (spoken/written language) or modality (verbal/auditory)	agnostic toward form and modality; open to tactile metaphors, etc.
continuous with other metaphor and with the “literal”	discrete; easily distinguishable from other metaphor and from the “literal”

## 6. Conclusions

This chapter began from the premise that GOFAI was never so “bad”FAI as it is often made out to be but that, nevertheless, like any field of inquiry, cognitive science is up for periodic renewal, and cognitive semiotics – with its focus on the meaning of meaning; with its prioritizing of dynamic processes over stable or static entities; with its inspiration from enactivism, phenomenology, and semiotics; with its inclination toward theoretical and



methodological pluralism – provides one highly promising opportunity to do so. A proper understanding of concepts is necessary for understanding both mind in general and consciousness in particular, where concepts and consciousness are two sides of one coin. However, too much of the thinking about concepts in “traditional” cognitive science remains mired in discussions about (generally symbolic – indeed generally propositional, explicit as opposed to implicit, etc.) knowledge representation.

Cognitive semiotics has much to inform theories of concepts regarding e.g. the nature of phenomenal experience and the *lifeworld* but, even more basically, to break discussion of concepts out of the traps into which it has all too frequently fallen: allowing one to see concepts as dynamic rather than fixed, prior to and foundational to language rather than part-and-parcel with, of one nature when we stop and reflect on them but logically of another when we possess and employ them non-reflectively, and so on. The hope is that a better understanding of concepts can, in turn, inform cognitive semiotics’ pursuit of a better understanding of the uniquely human place in the world. In the meantime, by working together, cognitive semiotics and theories of concepts can inform a better understanding of metaphor: one that is *not* based on antiquated notions of a strict and far-from-unproblematic distinction between literal and metaphorical.

### **Acknowledgments**

The author gratefully acknowledges the financial support and creative intellectual environment of the Centre for Cognitive Semiotics at Lund University, where much of the work for the present chapter was conducted; his colleagues in the Department of Cognitive Neuroscience and Philosophy at the University of Skövde; as well as useful feedback received during his presentation at the *International Association for Cognitive Semiotics* (IACS) conference in Lund, Sweden, in September 2014, and from the reviewers of earlier versions of this text.