Chapter 12

The symbol grounding problem has been solved, so what's next?

Luc Steels

12.1 Introduction

In the 1980s, a lot of ink was spent on the question of symbol grounding, largely triggered by Searle's Chinese room theory (Searle 1980). Searle's article had the advantage of stirring up discussion about when and how symbols could be about things in the world, whether intelligence involves representations or not, what embodiment means, and under what conditions cognition is embodied. But almost 25 years of philosophical discussion have shed little light on the issue, partly because the discussion has been mixed up with emotional arguments whether artificial intelligence (AI) is possible or not. However, today I believe that sufficient progress has been made in cognitive science and AI so that we can say that the symbol grounding problem has been solved. This chapter briefly discusses the issues of symbols, meanings, and embodiment (the main themes of the workshop), why I claim the symbol grounding problem has been solved, and what we should do next.

12.2 Symbols

As suggested in Chapter 2, let us start from Peirce and the (much longer) semiotic tradition which makes a distinction between a symbol, the objects in the world with which the symbol is associated (for example, for purposes of reference), and the concept associated with the symbol (see Figure 12.1). For example, we could have the symbol 'ball', a concrete round spherical object in the world with which a child is playing, and the concept *ball* which applies to this spherical object so that we can refer to the object using the symbol 'ball'.

In some cases, there is a method that constrains the use of a symbol for the objects with which it is associated. The method could, for example, be a classifier – a perceptual/pattern recognition process that operates over sensorimotor data to decide whether the object 'fits' with the concept. If such an effective method is available, then we call the symbol grounded. There are a lot of symbols which are not about the real world but about abstractions of various sorts, like the word 'serendipity', or about cultural meanings, like the word 'holy water', and so they will never be grounded through perceptual processes. In this chapter I focus on groundable symbols.

12.2.1 **Semiotic networks**

Together with the basic semiotic relations in Figure 12.1, there are some additional semantic relations that provide additional pathways for navigation between concepts, objects, and symbols.

- ◆ Objects occur in a context and may have other domain relationships with each other, for example, hierarchical or spatial and temporal relations.

- ◆ Symbols co-occur with other symbols in texts and speech, and this statistical structure can be picked up using statistical methods (as in the latent semantic analysis proposal put forward by Landauer and Dumais [1997]).

- ◆ Concepts may have semantic relations among each other, for example, because they tend to apply to the same objects.

- ◆ There are also relations between methods, for example, because they both use the same feature of the environment or use the same technique for classification.

Humans effortlessly use all these links and make jumps far beyond what would be logically warranted. This is nicely illustrated with social tagging, a new web technology whose usage exploded in recent years (Golder and Huberman 2006; Steels 2006). On sites like Flickr or last.fm, users can associate tags (i.e. symbols) with pictures or music files (i.e. objects). Some of these tags are clearly related to the content of the picture and hence could, in principle, be grounded using a particular method. For example, a picture containing a significant red object may be tagged 'red', a piano concerto may be tagged 'piano'. But most tags are purely associative; for example, a picture containing a dog may be tagged 'New York' because the picture was taken in New York.

Tagging sites compute and display the various semantic relations between tags and objects: they display the co-occurrence relations between tags so that users can navigate between the most widely used tags, as well as the contexts in which objects occur, for example, all pictures taken by the same user. The enormous success of these sites shows that this kind of system resonates strongly with large numbers of people, and I believe this is the case because it reflects and externalizes the same sort of semiotic relationships and navigation strategies that our brains are using in episodic memory or language.

I will call the huge set of links between objects, symbols, concepts, and their methods, a *semiotic network*. Every individual maintains such a network which is entirely his or her own.
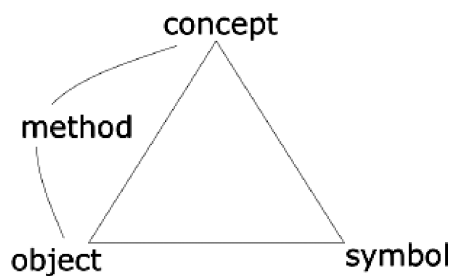

Fig. 12.1 The semiotic triad relates a symbol, an object, and a concept applicable to the object. The method is a procedure to decide whether the concept applies or not.

This semiotic network gets dynamically expanded and reshuffled every time you experience, think, or interact with the world or with others. Individuals navigate through this network for purposes of communication: when a speaker wants to draw the attention of an addressee to an object, he can use a concept whose method applies to the object, then choose the symbol associated with this concept and render it in speech or some other medium. The listener gets the symbol, uses his own vocabulary to retrieve the concept, and hence the method, and applies the method to decide which object might be intended.

Thus, if the speaker wants a bottle of wine on the table he is sitting at, he has to categorize this object as belonging to a certain class, then he has to look up the symbol associated with this class and use it to refer to the bottle. He might say, for example, 'Could you pass me the wine please?' Clearly, grounded symbols play a crucial role in communication about the real world, but the other links may also play important roles. For example, notice that the speaker said 'Could you pass me the wine please?,' whereas in fact he wanted the bottle that contained the wine. So the speaker navigated from container to content in his conceptualization of this situation, a move we commonly make. The speaker could also have said 'the Bordeaux please' and nobody would think that he requested the city of Bordeaux. Here, he navigated from container (bottle) to content (wine), from content (wine) to specific type of content (Bordeaux wine) characterized by the location where the content was made. Note that 'bottle of wine' might potentially be grounded in sensorimotor interaction with the world by some embodied pattern recognition method, but deciding whether a bottle contains Bordeaux wine is already a hopeless case, unless you can read and parse the label or are a highly expert connoisseur of wine.

There is apparently a debate in cognitive psychology between those emphasizing the grounded use of symbols (e.g. Barsalou 1999) and those emphasizing extraction and navigation across nongrounded semantic networks, and the current book reflects some of this debate. However, I think the opposition is a bit of a red herring. Both aspects of semantic processing are clearly extremely important and they interact with each other in normal human cognition. In this chapter I focus on grounded symbols not because I think other kinds of semantic processing are not relevant or important, but because in the Chinese room debate it is accepted that semantic processing can be done by computational systems whereas symbol grounding cannot.

12.2.2 Collective semiotic dynamics

Symbols can play a significant role in the development of an individual (see the example of children's drawings later in this chapter), but most of the time symbols are part of social interaction – such as in communication through language – and partners get feedback on how their own semiotic networks are similar or divergent from those of others. The semiotic networks that each individual builds up and maintains are therefore coupled to those of others and they get progressively coordinated in a group, based on feedback about their usage. If I ask for the wine and you give me the bottle of vinegar, both of us then learn that sometimes a bottle of vinegar looks like a bottle of wine. So we need to expand our methods for grounding 'wine' and 'vinegar' by tightening up the methods associated with the concepts that they express.

Psychological evidence for this progressive and continuous adaptation of semiotic networks is now beginning to come from many sources. First of all there are the studies of natural dialogue (Pickering and Garrod 2004; Clark and Brennan 1991) which show convincingly that speakers and hearers adopt and align their communication systems at all levels within the course of a single conversation. Their sound systems and gestures become similar, they adopt and negotiate new word meanings, they settle on certain grammatical constructions, and align their conceptualizations of the world. There are studies of emergent communication in laboratory conditions (Galantucci 2005) which show that the main skill required to establish a shared communication is the ability to detect miscommunication and repair them by introducing new symbols, changing the meaning of symbols, or adjusting your own conceptualizations based on taking the perspective of your partner. Growing evidence from linguistics (Francis and Michaelis 2005) show how human speakers and hearers engage in intense problem solving to repair inconsistencies or gaps in their grammars, and thus expand and align their linguistic systems with each other.

I call the set of all semiotic networks of a population of interacting individuals a *semiotic landscape.* Such a landscape is undergoing continuous change as every interaction may introduce, expand, or enforce certain relationships in the networks of individuals. Nevertheless there are general tendencies in the semiotic networks of a population, otherwise communication would not be possible. For example, it can be expected that individuals belonging to the same language community will have a similar network of concepts associated with the symbol 'red', and that some of these are grounded in a sufficiently similar way into the hue and brightness sensations so that if one says 'give me the red ball' she gets the red ball and not the blue one. Some tendencies also appear in other semantic relations between concepts. For example, in the Western world, the concept of 'red' is associated with danger, hence it is used in stop signs or traffic lights, whereas in some Asian cultures (e.g., China) 'red' is associated with joy.

Despite strong tendencies towards convergence, individual semiotic networks will never be exactly the same, even between two people interacting every day, as they are so much tied into personal histories and experiences. Psychological data confirms this enormous variation in human populations, even within the same language community (Webster and Kay 2005). The best one can hope for is that our semiotic networks are sufficiently coordinated to make joint action and communication possible.

12.2.3 **The symbol grounding problem**

Let me return now to the question originally posed by Searle (1980): can a robot deal with grounded symbols? More precisely, is it possible to build an artificial system that has a body, sensors and actuators, signal and image processing, pattern recognition processes, and information structures to store and use semiotic networks, and uses all that for communicating about the world or representing information about the world.

My first reaction is to say 'yes.' As far back as the early 1970s, AI experiments like Shakey the robot achieved this (Nilsson 1984). Shakey was a robot moving around and accepting commands to go to certain places in its environment. To act appropriately

upon a command like 'go to the next room and approach the big pyramid standing in the corner,' Shakey had to perceive the world, construct a world model, parse sentences and interpret them in terms of this world model, and then make a plan and execute it. Shakey could do all this. It was slow, but then it used a computer with roughly the same power as the processors we find in modern-day hotel doorknobs, and about as much memory as we find in yesterday's mobile phones.

So what was all the fuss created by Searle about? Was Searle's paper (and subsequent philosophical discussion) based on ignorance or on a lack of understanding of what was going on in these experiments? Probably partly. It has always been popular to bash AI because that puts one in the glorious position of defending humanity. But one part of the criticism was justified: all was programmed by human designers. The semiotic relations were not autonomously established by the artificial agent but carefully mapped out and then coded by human programmers. The semantics therefore came from us, humans. Nils Nilsson and the other designers of Shakey carved up the world, they thought deeply how the semantics of 'pyramid' and 'big' could be operationalized, and they programmed the mechanisms associating words and sentences with their meanings. So the Chinese room argument, if it is to make sense at all, needs to be taken differently, namely that computational system cannot generate their own semantics whereas natural systems (e.g., human brains) can. Indeed the mind/brain is capable to develop autonomously an enormous repertoire of concepts to deal with the environment and to associate them with symbols that are invented, adopted, and negotiated with others.

So the key question for symbol grounding is not whether a robot can be programmed to engage in some interaction which involves the use of symbols grounded in reality through his sensorimotor embodiment: that question has been solved. It is actually another question, well formulated by Harnad (1990): if someone claims that a robot can deal with grounded symbols, we expect that this robot autonomously establishes the semiotic networks that it is going to use to relate symbols with the world.

AI researchers had independently already come to the conclusion that autonomous grounding was necessary. By the late 1970s it was already clear that the methods needed to ground concepts, and hence symbols, had to be hugely complicated, domain-specific, and context-sensitive if they were to work at all. Continuing to program these methods by hand was therefore out of the question and that route was more or less abandoned. Instead, all effort was put into the machine learning of concepts, partly building further on the rapidly expanding field of pattern recognition, and partly by using 'neural network'-like structures such as perceptrons.

The main approach is based on supervised learning: the artificial learning system is shown examples and counterexamples of situations where a particular symbol is appropriate and it is assumed to learn progressively the grounding of the concepts that underlie the symbol in question, and hence to learn how to use the concept appropriately. We now have a wide range of systems and experiments demonstrating that this is entirely feasible. Probably the most impressive recent demonstration is in the work of Deb Roy and his collaborators (see Chapter 11, this volume). They provided example sentences and example situations to a vision-based robotic system and the robot was

shown to acquire progressively effective methods to use these symbols in subsequent real world interaction.

So does this mean that the symbol grounding problem is solved? I do not believe so. Even though these artificial systems now autonomously acquire their own methods for grounding concepts (and hence also symbols), it is still the human who sets up the world for the robot, carefully selects the examples and counterexamples, and supplies the symbol systems and conceptualizations of the world by drawing from an existing human language. So the semantics are still coming from us humans. Autonomous concept and symbol acquisition mimicking that of early child language acquisition is a very important step, but it cannot be the only one. The symbol grounding problem is not yet solved by taking this route and I believe it never will.

12.2.4 **Symbols in computer science**

Before continuing, I want to clear up a widespread confusion. The term 'symbol' has a venerable history in philosophy, linguistics, and other cognitive sciences, and I have tried to sketch its usage. However, it was hijacked in the late 1950s by computer scientists – more specifically AI researchers – and adopted in the context of programming language research. The notion of a symbol in so-called symbolic programming languages like LISP or Prolog is quite precise: it is a pointer (i.e., an address in computer memory) to a list structure containing a string known as the 'print name', which is used to read and write the symbol, possibly in addition to a value temporarily bound to this symbol, a definition of a function associated with this symbol, and an open-ended list of further properties and values associated with the symbol.

Thus, while the function of symbols can be recreated in other programming languages (like C++), the programmer then has to take care of allocating and deallocating memory for the internal pointer, reading and writing strings, and turning them into internal pointers, and he has to introduce more data structures for the other information items typically associated with symbols. In a symbolic programming language all that is done automatically. So a symbol is a very useful computational abstraction (like the notion of an array) and a typical LISP program might involve hundreds of thousands of symbols, created on the fly and reclaimed for memory ('garbage collected') when the need arises. Almost all sophisticated AI technology, as well as a lot of web technology, rests on the elegant but enormously powerful concept of symbolic programming.

Clearly this notion of symbol is not related to anything I discussed in the previous paragraphs, so I propose to make a distinction between *c-symbols* (the symbols of computer science)and *m-symbols* (meaning-oriented symbols in the tradition of the arts, humanities, and social and cognitive sciences). Unfortunately, when philosophers and cognitive scientists who are not knowledgable about computer programming read the AI literature they naturally apply the baggage of their own field, immediately assuming that if one uses a symbolic programming language, you must be talking about m-symbols and you possibly subscribe to all sorts of philosophical doctrines about symbols. Sadly, all this has given rise to what is probably the greatest terminological confusion in the history of science. The debate about the role of symbols in cognition or intelligence must

be totally decoupled to whether one uses a symbolic programming language or not. The rejection of 'traditional AI' by some cognitive scientists or philosophers seems mostly based on this misunderstanding. Thus, it is perfectly possible to implement a neural network using symbolic programming techniques, but these symbols are then c-symbols and not the fully-fledged 'semiotic' m-symbols I discussed earlier.

12.3 **Meaning**

The debate on symbol grounding is tied up with another discussion concerning the nature and importance of representations for cognition, and the relation between representations and meanings. The notion of representation also has a venerable history in philosophy, art history, etc., before it was hijacked by computer scientists to become much more narrow in scope. Since then, however, the computational notion of representation has returned to cognitive science through the back door of AI. Neuroscientists now talk without fear about representations and try to find their neural correlates. Let me try to trace these shifts, starting with the original precomputational notion of representation, as we find for example in the famous essay on the hobby horse by Gombrich (1969) or in the more recent writings of Bruner (1990).

12.3.1 **Representations and meanings**

In traditional usage, a representation is a stand-in for something else, so that it can be made present again (i.e. re-present-ed). Anything can be a representation of anything. For example, a pen can be a representation of a boat, a person, or an upward movement; a broomstrick can be a representation of a hobby horse. The magic of representations happens because one person decides to establish that *x* is a representation for *y*, and others either agree with this or accept that this representational relation holds. There is nothing in the nature of an object that makes it a representation or not; it is rather the role the object plays in subsequent interaction. Of course, it helps a lot if the representation has some properties that help others to guess what it might be a representation of. But a representation is seldom a picture-like copy of what is represented.

It is clear that m-symbols are a particular type of representations, one where the physical object being used has a rather arbitrary relation to what is represented. It follows that all the remarks made earlier about m-symbols are also valid for representations. However, the notion of representation is broader. Following Peirce, we can make a distinction between an icon, an index, and a symbol. An icon looks like the thing it signifies, so meaning arises by perceptual processes and assocations, the same way we perceive the objects themselves. An example of an icon is a statue of a saint which looks like the saint, or how people imagine the saint to be. An index does not look like the thing it signifies, but there is nevertheless some causal or associative relation. For example, smoke is an index of fire. A symbol on the other hand has no direct or indirect relation to the thing it signifies. The meaning of a symbol is established purely by convention, and hence you have to know the convention in order to figure out what the symbol is about.

Human representations have some additional features. First of all they seldom represent physical things, but rather meanings. A meaning is a feature that is relevant in the interaction between the person and the thing being represented. For example, a child may represent a fire engine by a square painted red. Why red? Because this is a distinctive important feature of the fire engine for the child. There is a tendency to confuse meanings and representations. A representation 're-presents' meaning but should not be equated with meaning, just like an ambassador may represent a country but is not equal to that country. Something is meaningful if it is important in one way or another: for survival, maintaining a job, social relations, navigating in the world, etc. For example, the differences in colour between different mushrooms may be relevant to me because they help to distinguish those that are poisonous from those that are not. If colour is irrelevant and shape instead is distinctive, then the shape features of mushroom would be meaningful (Cangelosi *et al.* 2000).

Conceiving a representation requires selecting a number of relevant meanings and deciding how these meanings can be invoked in ourselves or others. The invocation process is always indirect and most often requires inference. Without knowing the context, personal history, prior use of representations, etc., it is often very difficult, if not impossible, to decipher representations. Second, human representations typically involve perspective. Things are seen and represented from particular points of view and these perspectives are intermixed.

A nice example of a representation is shown in Figure 12.2. Perhaps you thought that this drawing represents a garden, with the person in the middle watering the plants. Some people report that they interpret this as a kitchen with pots and pans and somebody cooking a meal. As a matter of fact, the drawing represents a British double-decker bus (the word 'baz' is written on the drawing). Once you adopt this interpretation, it is easy to guess that the bus driver must be sitting on the right-hand side in his enclosure.

Fig. 12.2 Drawing by a child called Monica. Can you guess what meanings this representation expresses?

The conductor is shown in the middle of the picture. The top shows a set of windows of increasing size, and the bottom a set of wheels. Everything is enclosed in a big box which covers the whole page. This drawing was made by a child, Monica, aged 4.2 years.

Clearly this drawing is not a realistic depiction of a bus. Instead it expresses some of the important meanings of a bus from the viewpoint of Monica. Some of the meanings have to do with recognizing objects, in this case, recognizing whether a bus is approaching so that you can get ready to get on it. A bus is huge. That is perhaps why the drawing fills the whole page, and why the windows at the top and the wheels at the bottom are drawn as far apart as possible. A bus has many more windows than an ordinary car, so Monica has drawn many windows, and the same thing for the wheels: a bus has many more wheels than an ordinary car so many wheels are drawn. The exact number of wheels or windows does not matter, there must only be enough of them to express the concept of 'many'. The shape of the windows and the wheels has been chosen by analogy with their shape in the real world. They are positioned at the top and the bottom as in a normal bus viewed sideways. The concept of 'many' is expressed in the visual grammar invented by the child.

Showing your ticket to the conductor or buying a ticket must have been an important event for Monica. It is of such importance that the conductor, who plays a central role in this interaction, is put in the middle of the picture and drawn prominently to make them stand out as foreground. Once again, features of the conductor have been selected that are meaningful, in the sense of meaningful for recognizing the conductor. The human figure is schematically represented by drawing essential body parts (head, torso, arms, legs). Nobody fails to recognize that this is a human figure, which cannot be said for the way the driver has been drawn. The conductor carries a ticketing machine. Monica's mother, with whom I corresponded about this picture, wrote that this machine makes a loud 'ping' and would be impossible not to notice. There is also something drawn on the right side of the head, which is most probably a hat, another characteristic feature of the conductor. The activity of the conductor is represented, too: the right arm is extended as if ready to accept money or check a ticket.

Composite objects are superimposed and there is no hesitation to mix different perspectives. This happens quite often in children's drawings. If they draw a table viewed from the side, they typically draw all the forks, knives, and plates as separate objects 'floating above' the table. In Figure 12.2, a bird's eye perspective is adopted so that the driver is located on the right-hand side in a separate box and the conductor in the middle. At the same time, the bus is viewed from the side so that the wheels are at the bottom and the windows near the top. Then there is the third perspective of a sideways view (as if seated inside the bus), which is used to draw the conductor as a standing up figure. Even for drawing the conductor, two perspectives are adopted: the sideways view for the figure itself and the bird's eye view for the ticketing machine so that we can see what is inside. Multiple perspectives in a single drawing are later abandoned as children try to make their drawings more 'realistic', and hence less creative. But it may reappear again in artists' drawings. For example, many of Picasso's paintings play around with different perspectives on the same object in the same image.

The fact that the windows get bigger towards the front expresses another aspect of a bus which is most probably very important to Monica. At the front of a double-decker bus there is a huge window and it is great fun for a child to sit and watch the streets go by. The increasing size of the windows reflects the desirability to be near the front. Size is a general representational tool used by many children (and also in mediaeval art) to emphasize what is considered to be important. Most representations express not only facts but above all attitudes and interpretations of facts. Though the representations of very young children may seem to be random marks on a piece of paper, for them they are almost purely emotional expressions of attitudes and feelings, like anger or love.

Human representations are clearly incredibly rich and serve multiple purposes. They help us to engage with the world and share world views and feelings with others. Symbols or symbolic representations traditionally have this rich connotation. And I believe the symbol grounding problem can only be said to be solved if we understand, at least in principle, how individuals originate and choose the meanings that they find worthwhile to use as basis for their (symbolic) representations, how perspective may arise, and how the expression of different meanings can be combined to create compositional representations.

12.3.2 Representations in computer science

In the 1950s, when higher-level computer programming started to develop, computer scientists began to adopt the term 'representation' for data structures that held information for an ongoing computational process. For example, in order to calculate the amount of water left in a tank with a hole in it, you have to represent the initial amount of water, the flow rate, the amount of water after a certain time, etc., so that calculations can be done over these representations. Information processing came to be understood as designing representations and orchestrating the processes for the creation and manipulation of these representations.

Here, computer scientists (and *ipso facto* AI researchers) are clearly adopting only one aspect of representations: the creation of a 'stand-in' for something into the physical world of the computer so that it could be transformed, stored, transmitted, etc. They are not trying to operationalize the much more complex process of meaning selection, representational choice, composition, perspective taking, inferential interpretation, etc., all of which are an important part of human representation making. I will use the term *c-representations* to mean representations as used in computer science and *m-representations* for the original, meaning-oriented use of representations in social science and humanities.

In building AI systems there can be no doubt that c-representations are needed to do anything at all. It is also now accepted among computational neuroscience researchers that c-representations must be used in the brain (i.e., information structures for visual processing, motor control, planning, language, memory, etc.). Mirror neuron networks are a clear example of c-representations. If the same circuits become active both in the perception of an action and the execution of it, it means that these circuits act as representations of actions, usable as such by other neural circuits. So the question

today is no longer whether or not c-representations are necessary but rather what their nature might be.

As AI research progressed, a debate arose between advocates of 'symbolic' c-representations and 'nonsymbolic' c-representations. Simplified, symbolic c-representations are representations for categories, classes, individual objects, events, or anything else that is relevant for reasoning, planning, language, or memory. Nonsymbolic c-representations, on the other hand, use continuous values and are therefore much closer to the sensory and motor streams. They have also been called analogue c-representations because their values change by analogy with a state of the world, like a temperature measure which goes up and down when it is hotter or colder. Thus, we could have on the one hand a sensory channel for the wavelength of light with numerical values (a nonsymbolic c-representation), and on the other hand (symbolic) c-representations for colour categories like 'red', 'green', etc. Similarly, we could have an infrared or sonar channel that numerically reflects the distance of the robot to obstacles (a nonsymbolic c-representation), or we could have a c-symbol representing 'obstacle seen' (a symbolic c-representation).

Sometimes the term 'subsymbolic' c-representation is used to mean either nonsymbolic c-representation or a distributed symbolic c-representation, as employed in connectionist systems (Rumelhart and McClelland 1986; Smolensky 1988). The latter type assumes that the representation (i.e., the physical object representing a concept) is actually a cloud of more or less pertinent features, themselves represented as primitive nodes in the network.

When the neural network movement became more prominent again in the 1980s, they de-emphasized symbolic c-representations in favour of the propagation of (continuous) activation values in neural networks (although some researchers used these networks later in an entirely symbolic way, as pointed out in the discussion of Rogers *et al.* in Chapter 2). When Brooks wrote a paper on 'Intelligence without representation' (Brooks 1991), he argued that real-time robotic behaviour could often be better achieved without symbolic c-representations. For example, instead of having a rule that says 'if obstacle then move back and turn away from obstacle,' we could have a dynamical system that directly couples the change in sensory values to a change on the actuators. The more infrared reflection from obstacles is picked up by the right infrared sensor on a robot, the slower its left motor is made to move, and hence the robot starts to veer away from the obstacle.

But does this mean that all symbolic c-representations have to be rejected? This does not seem to be warranted, particularly not if we are considering language processing, expert problem solving, conceptual memory, etc. At least from a practical point of view it makes much more sense to design and implement such systems using symbolic c-representations and mix symbolic, nonsymbolic, and subsymbolic c-representations whenever appropriate. The debate between symbolic or nonsymbolic representations seems to be based on an unnecessary opposition, just as the debate between grounded and non-grounded symbols. The truth is that we need both.

At the same time, the simple fact of using a c-representation (symbolic or otherwise) does not yet mean that an artificial system is able to come up or interpret the meanings

that are represented by the representation. Meaning and representation are different things. In order to see the emergence of meaning, we need a minimum of a task, an environment, and an interaction between the agent and the environment that works towards an achievement of the task.

12.4 **Embodiment**

We have seen that the notions of symbol and representation are used quite differently by computer scientists than by cognitive scientists, muddling the debate about symbol grounding. Computer science symbols and representations capture only a highly limited aspect of what social and cognitive scientists mean by symbols and representations. This has also happened, I think, with a third notion the debate has considered: embodiment.

12.4.1 **Embodiment as implementation**

In the technical literature (e.g., in patents), the term embodiment refers to implementation (i.e., the physical realization) of a method or idea. Computer scientists more commonly use the word *implementation*. An implementation of an algorithm is a definition of that algorithm in a form such that it can be physically instantiated on a computer and actually run (i.e., that the various steps of the algorithm find their analogue in physical operations). It is absolute dogma in computer science that an algorithm can be implemented in many different physical media, on many different types of computer architectures, and in different programming languages, even though, of course, it may take much longer (either to implement or to execute) in one medium versus another.

Computer scientists naturally take a systems perspective when they try to understand a phenomenon. In contrast, most physical natural scientists seek material explanations, they seek an explanation in the materials and material properties involved in a phenomenon. For example, their explanation of why oil floats on top of water is in terms of the attraction properties of molecules, the density of oil and water, and the upward buoyancy forces. System explanations, on the other hand, are in terms of elements and processes between the elements. For example, the explanation for money is not in terms of the material properties of coins or bank notes, but in terms of legal agreements, conventions, central banks, trust relations, etc. The franc, lire, or Deutschmark were replaced overnight by the Euro and everything kept going.

It is therefore natural that computer scientists apply system thinking to cognition and the brain. Their goal is to identify the algorithms (in other words the systems) underlying cognitive activities like symbol grounding and study them through various kinds of implementations which may not at all be brain-like. They might also talk about neural implementation or neural embodiment, meaning the physical instantiation of a particular algorithm with the hardware components and processes available to the brain. But most computer scientists know that the distance between an algorithm (even a simple one) and its embodiment, for example in an electronic circuit, is huge, with many layers of complexity intervening. Moreover, the translation through all these layers is never done by hand but by compilers and assemblers. It is very possible that we will have to

follow the same strategy to bridge the gap between high-level models of cognition and low-level neural implementations.

There has been considerable resistance both from biologists and philosophers alike for accepting the systems perspective, i.e., the idea that the same process (or algorithm) can be instantiated (embodied) in many different media. They seem to believe that the (bio-)physics of the brain must have unique characteristics with unique causal powers. For example, Penrose (1989) has argued that intelligence and consciousness is due to certain (unknown) quantum processes which are unique to the brain, and therefore any other type of implementation can never obtain the same functionality. This turns out to be also the fundamental criticism of Searle. The reason why he argues that artificial systems, even if they are housed in robotic bodies, cannot deal with symbols is because they will for ever lack *intentionality*. Intentionality is 'that feature of certain mental states by which they are directed at or about objects and states of affairs in the world' (Searle 1980, footnote 2), and it is of course essential for generating and using symbols about the world. According to Searle, an adequate explanation for intentionality can only be a material one: 'Whatever else intentionality is, it is a biological phenomenon, and it is as likely to be as causally dependent on the specific biochemistry of its origins as lactation, photosynthesis, or any other biological phenomena.' If that is indeed true, investigating the symbol grounding problem through experiments with artificial robotic agents is totally futile.

Searle invokes biology, but most biologists today accept that the critical features of living systems point to the need to adopt a system perspective instead of a material one; even photosynthesis can be done through many different materials. For example, leading evolutionary biologist John Maynard Smith (2000) has been arguing clearly that the genetic system is best understood in information processing terms and not (just) in molecular terms. Neurobiologists Gerald Edelman and Giulilo Tononi (2000) have proposed system explanations for phenomena-like consciousness in terms of re-entrant processing and coordination of networks of neural maps, rather than specific biosubstances or quantum processes. Thus, the system viewpoint is gaining more and more prominence in biology rather than diminishing in productivity.

12.4.2 Embodiment as having a physical body

There is another notion of embodiment that is also relevant in this discussion. This refers quite literally to 'having a body' for interacting with the world, i.e. having a physical structure, dotted with sensors and actuators, and the necessary signal processing and pattern recognition to bridge the gap from reality to symbol use. Embodiment in this sense is a clear precondition to symbol grounding. As soon as we step outside the realm of computer simulations and embed computers in physical robotic systems, we achieve some form of embodiment (even if the embodiment is not the same as human embodiment). It is entirely possible, and in fact quite likely, that the human body and its biophysical properties for interacting with the world make certain behaviours possible and allow certain forms of interaction that are unique. This puts a limit on how

similar AI systems can be to human intelligence. But it only means that embodied AI is necessarily of a different nature due to the differences in embodiment, not that embodying AI is impossible.

12.5 **A solution to the symbol grounding problem?**

Over the past decade I have been working with a team of a dozen graduate students at the University of Brussels (VUB AI Lab) and the Sony Computer Science Lab in Paris on various experiments in language emergence (Steels 2003). They take place in the context of a broader field of study concerned with modelling language evolution (see Minett and Wang 2005 or Vogt 2006 for recent overviews; see Vogt 2002 for other examples). We are carrying them out to investigate many aspects of language and cognition, but here I focus on their relevance for the symbol grounding problem. I will illustrate the ingredients that we put in these experiments using a specific example of a colour guessing game, discussed in much more detail in Steels and Belpaeme (2005).

12.5.1 **Embodiment**

A first prerequisite for solving the symbol grounding problem is that we can work with physically embodied autonomous agents, autonomous in the sense that they have their own source of energy and computing power, they are physically present in the world through a body, they have a variety of sensors and actuators to interact with the world, and, most importantly, they move and behave without any remote control or further human intervention once the experiment starts. All these conditions are satisfied in our experiments.

We have been using progressively more complex embodiments, starting from pan-tilt cameras in our earlier 'Talking Heads' experiments, moving to Sony AIBO dog-like robots, and, more recently, fully-fledged humanoid QRIO robots (see Figure 12.2). These robots are among the most complex physical robots currently available. They have a humanoid shape with two legs, a head, and two arms with hands and fingers. The robots are 0.6 metres (2 feet) tall and weigh 7.3 kilograms (16 pounds). They have cameras for visual input, microphones for audio input, touch and infrared sensors, and a large collection of motors at various joints, with sensors at each motor. The robots have a huge amount of computing power on board with general- and special-purpose processors. In addition they are wirelessly connected to off-board computers which can be harnassed to increase computing power, up to the level of supercomputers. Even extremely computation-intensive image processing and motor control is possible in real time. The robots have been programmed using a behaviour-based approach (Brooks 1991) so that obstacle avoidance, locomotion, tracking, grasping, etc., are all available as solid smooth behaviours to build upon.

By using these robots, we achieve the first prerequisite for embodied cognition and symbol grounding, namely that there is a rich embodiment. Often we use the same robot bodies for different agents by uploading and downloading the complete state of an agent after and before a game. This way we can do experiments with larger population sizes.

12.5.2 **Sources of meaning**

If we want to solve the symbol grounding problem, we next need a mechanism by which an (artificial) agent can autonomously generate its own meanings. This means that there must be distinctions that are relevant to the agent in its agent–environment interaction. The agent must therefore have a way to introduce new distinctions based on the needs of the task. As also argued by Cangelosi *et al.* (2000), this implies that there must be a task setting in which some distinctions become relevant and others do not.

One could imagine a variety of activities that generate meaning, but we have focused on language games. A language game is a routinized situated interaction between two embodied agents who have a cooperative goal (e.g., one agent wants to draw the attention of another agent to an object in the environment, or one agent wants the other one to perform a particular action) and who use some form of symbolic interaction to achieve that goal (e.g., by exchanging language-like symbols or sentences, augmented with nonverbal interactions such as pointing or bodily movement towards objects).

In order to play a language game the agents need a script with which they can establish a joint attention frame, in the sense that the context becomes restricted and it is possible to guess meaning and interpret feedback about the outcome of a game. Specifically, in the colour guessing game shown in Plate 12.1 (the so-called Mondriaan experiment), agents walk towards a table on which there are colour samples. One agent randomly becomes speaker, chooses one sample as topic, uses his available grounding methods to categorize the colour of this sample as distinctive from the other colour samples, and names the category. The other agent is hearer; it decodes the name to retrieve the colour category, uses the method associated with this category to see to which sample it applies, and then points to the sample. The game is a success if the speaker agrees that the hearer pointed to the sample it had originally chosen as topic.

Needless to say that the robots have no idea which colours they are going to encounter. We introduce different samples and can therefore manipulate the environment driving the categories and symbols that the agents need. For example, we can only introduce samples with different shades of blue (as in Plate 12.1) which will lead to very fine-grained distinctions in the blue range of the spectrum, or we can spread the colours far apart.

12.5.3 **Grounding of categories**

Next we need a mechanism by which agents can internally represent and ground their relevant meanings. In the experiment, agents start with no prior inventory of categories and no inventory of methods (classifiers) that apply categories to the features (sensory experience) they extracted from the visual sensation they received through their cameras. In the experiments, the classifiers use a prototype-based approach implemented with radial basis function networks (Steels and Belpaeme 2005). A category is distinctive for a chosen topic if the colour of the topic falls within the region around a particular prototype and all other samples fall outside of it. For example, if there is a red, green, and blue sample and the red one is the topic, then if the topic's colour falls within the region

around the red prototype and the others do not, red is a valid distinctive category for the topic. If an agent cannot make a distinction between the colour of the topic and the colour of the other samples in the context, it introduces a new prototype and will later progressively adjust or tighten the boundaries of the region around the prototype by changing weights.

12.5.4 Self-organization of symbols

The next requirement is that agents autonomously can establish and negotiate symbols to express the meanings that they need to express. In the experiment, agents generate new symbols by combining randomly a number of syllables into a word, like 'wabado' or 'bolima'. The meaning of a word is a perceptually grounded category. No prior lexicon is given to the agents, and there is no central control that will determine by remote control how each agent has to use a word. Instead, a speaker invents a new word when it does not have a word yet to name a particular category; the a hearer will try to guess the meaning of the unknown word based on feedback after a failed game, and thus new words enter into the lexicons of the agents and propagate through the group.

Coordination process

If every agent generates his own meanings, perceptually grounded categories, and symbols, then no communication is possible, so we need a process of coordination that creates the right kind of semiotic dynamics so that the semiotic networks of the individual agents become sufficiently coordinated to form a relatively organized semiotic landscape.

This is achieved in two ways. Firstly, speakers and hearers continue to adjust the score of form-meaning associations in their lexicon based on the outcome of a game: When the game is a success they increase the score and dampen the score of competing associations; when the game is a failure the score is diminished. The net effect of this update mechanism is that a positive feedback arises: words that are successful are used more often and hence become even more successful. After a while the population settles on a shared lexicon (see Figure 12.3).

Speakers and hearers also maintain scores about the success of perceptually grounded categories in the language game, and adjust these scores based on the outcome. As a consequence, the perceptually grounded categories also get coordinated in the sense that they become more similar, even though they will never be absolutely identical. This is shown for the colour guessing game in Plate 12.2. We note that if agents just play discrimination games but do not use their perceptually grounded categories as part of language games, they succeed in discrimination but their categorical repertoires will show much more variation.

We have been carrying out many more experiments addressing issues related to the origins of more complex meanings, the more complex use of embodiments (e.g., in action), and the emergence of more complex human language-like symbolic representations, even with grammatical structure (Steels 2005). For example, our 'perspective reversal experiment' has demonstrated how agents self-organize an inventory of spatial categories and spatial language symbolizing these categories, as well as the ability to use perspective reversal and the marking of perspective (Steels and Loetzsch 2007).
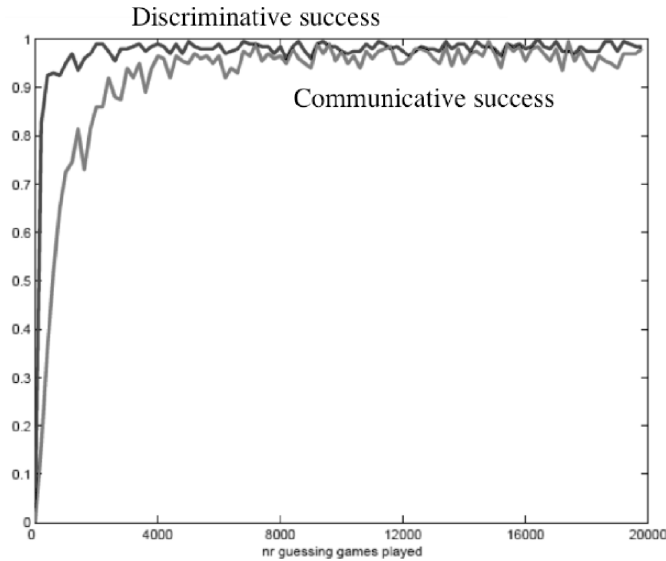
Fig. 12.3 Results of experiments showing the coordination of colour categories in a cultural evolution process. A population of 20 agents play guessing games naming colour distinctions and coordinating both their lexicon and colour categories through the game. We see progressively that discriminative success and communicative success (y-axis) increase as the number of games progresses (x-axis).

12.5.5 **So, have we solved the grounding problem?**

I argue that these experiments show that we have an effective solution to the symbol grounding problem, if there is ever going to be one: we have identified the right mechanisms and interaction patterns so that the agents autonomously generate meaning, autonomously ground meaning in the world through a sensorimotor embodiment and perceptually grounded categorization methods, and autonomously introduce and negotiate symbols for invoking these meanings. The objective test for this claim is in the increased success of agents in the language games. Clearly, if the agents do not manage to generate meanings and coordinate their perceptually grounded categories and symbols, they will have only a random chance of succeeding, whereas we see that they reach almost total success in the game. There is no human prior design to supply the symbols or their semantics, neither by direct programming nor by supervised learning.

The explanatory power of these experiments does not come from the identification of some biochemical substance for intentionality of the sort Searle is looking for, but it is a system explanation in terms of semiotic networks and semiotic dynamics operating over these networks. Each agent builds up a semiotic network relating sensations and sensory experiences to perceptually grounded categories and symbols for these categories (see Plate 12.3). All the links in these networks have a continuously valued strength or score that is continually adjusted as part of the interactions of the agent with the environment and with other agents. Links in the network may at any time be added

or changed as a side effect of a game. Although each agent does this locally, an overall coordinated semiotic landscape arises.

12.6 **Discussion**

The main goal of this chapter was to clarify some terminology and attempt to indicate where we are with respect to some of the most fundamental questions in cognition, namely the symbol grounding problem. From the viewpoint of AI, the question is whether it is possible to ever conceive of an artificial system that is able to invent and use grounded symbols in its sensorimotor interactions with the world and others. Several discussants, most notably Searle, have argued that this will never be possible because artificial systems will forever lack the critical biomaterial substance. I now boldly state that the symbol grounding problem is solved, by that I mean we now understand enough to create experiments in which groups of agents self-organize symbolic system that are grounded in their interactions with the world and others.

So where do we go from here? Clearly there is still a lot to learn. First of all, we can do many more experiments with artificial robotic agents to progressively understand many more aspects of meaning, conceptualization, symbolization, and the dynamical interactions between them. These experiments might focus, for example, on issues of time and its conceptualization in terms of tense, mood, modality, and the roles of objects in events, as well as its expression in case grammars, the categorization of physical actions and their expression, raising issues in the mirror neuron debate, etc. I see this as the job of AI researchers and (computational) linguists.

Second, we can investigate whether there are any neural correlates in the brain for the semiotic networks being discussed here, which would be impossible to find with local cell recordings but would require looking at global connections and long-distance correlations between neural group firings (Sporns and Tononi 2005). The work of Pulvermüller (Chapter 6, this volume) has already been able to identify some of these networks and shown that they span the whole brain. For example, smell words activate areas for olfactory processing or action words activate cells in motor circuitry. We can also investigate whether we can find neural correlates for the dynamical mechanisms that allow individuals to participate in the semiotic dynamics going on in their community. All this is going to be the job of cognitive neuroscientists.

Finally, there is a need for new types of psychological observations and experiments investigating representation-making in action (for example in dialogue or drawing) and investigating group dynamics. The experiments I have been discussing provide models to understand how humans manage to self-organize grounded communication systems, although I have produced no evidence in this chapter for such a claim: this is the job of the experimental psychologists.

All too often psychological experiments have focused on the single individual in isolated circumstances, whereas enormously exciting new discoveries can be expected if we track the semiotic dynamics in populations, something that is becoming more feasible thanks to internet technologies. Psychological experiments have often assumed as

well that symbol use and categorization is static, whereas the interesting feature is precisely in its dynamics.

I believe there has been clear progress on the issue of symbol grounding and that there will be much more progress in the coming decade, provided we keep an open mind, engage in interdisciplinary curiosity, and avoid false debates.

Debate

Arthur Graesser: I was curious whether you tried to have your robots handle 'my,' 'your,' or 'our' with a modifier, which would be very diagnostic of perspective and point of view, like 'on *my* left,' 'on *your* left,' 'on *the* left,' 'on *our* left,' and things like that. And that's where you also get syntax coming in, and ordering. Is there any emergent way that was helping them figure out these different points of view?

Luc Steels: Well, right now in the experiment there is just my point of view and your point of view. But we are currently doing one in which there would be several points of view, like a third robot who also sees the situation. We needed to set up a communicative challenge to see the thing emerge. I think syntax is not necessarily because of that. You know, we can go into that, but that's a bigger topic. I have precise ideas about this, which I'm willing to share, but maybe not now.

Mitchell Nathan: So Luc, in one of your very first slides, you made this pure convention claim, but what if we did the following experiment, if we ask people to decide which of the labels 'tanak' or 'oblio' should be assigned to a sawtooth image and which to a smoothly curving image? Would it come out as purely arbitrary or would we find some agreement, and if we did, what does that imply?

Steels: Well we could do the experiment later, maybe during the coffee break. But personally, I think the assignment would be arbitrary. If it's really names.

Deb Roy: I had a question. I noticed several times you emphasized that things were being learnt, that they weren't innate. But I thought you are modelling evolutionary processes, so I was wondering if you could clarify. Because typically when people make that distinction of learnt versus innate, they mean these are not evolved structures; these are lifetime learnt. And yet all of the experiments, as I understood them, were models of evolutionary processes. So are you seeing a relationship here between what we call language acquisition in a lifetime versus evolution of language, or are you treating this process as sort of contiguous of both? What do you mean by innate versus learnt?

Steels: This is a very good point, which I didn't clarify. So, these things are cultural processes, it's not evolution over centuries. I mean this happens very quickly, right? And I think what psychological dialogue studies have shown us in the last 10 years is in fact that human beings, when they sit together and go into a dialogue, they continuously adapt their communication system at all levels. And this is going on here, too. Now, the model we should adopt for understanding symbolization is not that at one time somebody symbolized. It's a continuously adaptive process, you know, all the time changing. It's true that here we start from scratch – this is because I wanted to

emphasize that we didn't put anything in, and also to find an explanation for the origin of these things; how it is possible that language has emerged? But in fact if there's already a system, then you can throw in new agents and they will learn preferably the system that's already there. This is because they are exposed with a higher frequency to existing conventions and so they absorb them. But, in that sense, it's a cultural process. Also, I briefly mentioned recruitment. This is obviously happening in developmental time, so when you pull in new systems – an emotional system, perspective transform, or what have you – they are all pulled into the language faculty in developmental time. So in terms of what is innate I would say it's the fundamental hardware of all these mechanisms, like bidirectional associative memory, etc., dealing with sequences – all of that stuff. You have to pull them in to be able to participate in the collective construction of language.

Roy: We talked earlier about the process of symbolization, building on things you hinted at. I was wondering what might be uniquely human about that process. Sort of the possibility of looking for what is maximally distinctive between two things and having the motor repertoire to mimic it, or to indicate it.

Lawrence Shapiro: I was wondering what role embodiment plays in the learning of these categories. It occurred to me that it depends on the sort of categories that you want to learn. We could play this game just sitting around the table here and that doesn't seem to require embodiment in the senses that I've heard others discussing.

Steels: Well, if you play around the table we need still the perceptual system, we need the pointing for gestures. But this is, I think, where this perspective reversal experiment is relevant. Because Chomsky, for example, would say language has nothing to do with communication, right? It evolved for internal representation and then, accidentally, it became used externally through some translation. Now then, how would you explain that perspective marking is so common in languages, right? 'Your left,' 'my left,' etc. So, the problem of perspective is a direct consequence of the problem of being embodied in the world. We had to put that in these different, unpredictable perspectives to get this feature of language going. That's part of the answer.

Friedemann Pulvermüller: Thanks for an excellent talk, very interesting. I'm not sure whether I fully followed and so allow me a potentially very stupid question. What does it buy the agents if they align their word usage?

Steels: Well, if I say 'ba' and this means left for you but right for me, or move forward for you and move left for me, we're going to have failure in communication, right? The whole process is tremendously difficult because every situation can be conceptualized from many points of view. You get a search process. You're dealing with noisy sensing, you know, multiple world models. So you're swimming in a sea of uncertainty and you try to constrain the degrees of freedom that you need to explore. Now the more you are aligned, the higher your chances are to succeed. And so that's why I think you see in many experiments now, of Simon Garrod and others, that people engaged in dialogue on the spot very quickly invent new words and new grammatical constructions.

Author note

References

Barsalou L (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609.

Brooks R (1991). Intelligence without representation. *Artificial Intelligence Journal*, 47, 139–59.

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

Cangelosi A, Greco A, Harnad S (2002). Symbol grounding and the symbolic theft hypothesis. In A Cangelosi, D Parisi, Eds. (2000). *Simulating the Evolution of Language.* Berlin: Springer Verlag.

Clark H, Brennan S (1991). Grounding in communication. In: L Resnick, S Levine, S Teasley, Eds. *Perspectives on Socially Shared Cognition* (pp. 127–49). Washington, DC: APA Books.

Davidoff J (2001). Language and perceptual categorisation. *Trends in Cognitive Sciences*, 5, 382–7.

Edelman G (1999). *Bright Air, Brilliant Fire: On the Matter of the Mind.* New York, NY: Basic Books.

Edelman G, Tononi G (2000). *A Universe of Consciousness. How Matter Becomes Imagination.* New York, NY: Basic Books.

Francis E, Michaelis L (2002). *Mismatch: Form-Function Incongruity and the Architecture of Grammar.* Stanford, CA: CSLI Publications.

Galantucci B (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29, 737–67.

Golder S, B Huberman (2006). The structure of collaborative tagging. *Journal of Information Science*, 32, 198–208.

Gombrich EH (1969). *Art and Illusion: A Study in the Psychology of Pictorial Representation.* Princeton, NJ: Princeton University Press.

Harnad S (1990). The symbol grounding problem. *Physica D*, 42, 335–46.

Kay P, Regier T (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences USA*, 100, 9085–9.

Landauer TK, ST Dumais (1997). A solution to Plato's problem: the latent semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104, 211–40.

Maynard Smith J (2000). The concept of information in biology. *Philosophy of Science*, 67, 177–94.

Minett JW, Wang WS-Y (2005). *Language Acquisition, Change and Emergence: Essays in Evolutionary Linguistics.* Hong Kong: City University of Hong Kong Press.

Nilsson NJ (1984). Shakey the Robot, Technical Note 323. Menlo Park, CA: AI Center, SRI International.

Penrose R (1989). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics.* Oxford: Oxford University Press.

Pickering MJ, Garrod S (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–225.

Rumelhart DE, McClelland JL; PDP Research Group (1986). *Parallel Distributed Processing.* Cambridge, MA: MIT Press.

Searle JR (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3, 417–57

Smolensky P (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11, 1–74.

Sporns O, Tononi G, Kotter R (2005). The human connectome: a structural description of the human brain. *PLoS Computational Biology*, 1, e42.

Steels L (2003). Evolving grounded communication for robots. *Trends in Cognitive Science*, 7, 308–12.

Steels L (2004). Intelligence with representation. *Philosophical Transactions of the Royal Society of London A*, 361, 2381–95.

Steels L (2005). The emergence and evolution of linguistic structure: from minimal to grammatical communication systems. *Connection Science*, 17, 213–30.

Steels L (2006). Collaborative tagging as distributed cognition. *Pragmatics and Cognition*, 14, 287–92.

Steels L, Belpaeme T (2005). Coordinating perceptually grounded categories through language: a case study for colour. *Behavioral and Brain Sciences*, 24, 469–89.

Steels L, Loetzsch M (2007). Perspective alignment in spatial language. In KR Coventry, T Tenbrink, JA Bateman, Eds. *Spatial Language and Dialogue*. Oxford: Oxford University Press.

Vogt P (2002). Physical symbol grounding. *Cognitive Systems Research*, 3, 429–57.

Vogt P, Sugita Y, Tuci E, Nehaniv C, Eds (2006). *Symbol Grounding and Beyond*. Berlin: Springer Verlag.

Webster MA, Kay P (2005). Variations in color naming within and across populations. *Behavioral and Brain Sciences*, 28, 512–13.